

Form PTO-1390  U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE  TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		ATTORNEY'S DOCKET NUMBER P20934 <b>09/807427</b> U.S. APPLICATION NO. (If known, see 37 CFR 1.5)
INTERNATIONAL APPLICATION NO. PCT/JP00/05621	INTERNATIONAL FILING DATE 23 August 2000	PRIORITY DATE CLAIMED 23 August 1999
TITLE OF INVENTION APPARATUS AND METHOD FOR SPEECH CODING		
APPLICANT(S) FOR DO/EO/US Kazutoshi YASUNAGA and Toshiyuki MORII		
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information.		
1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.		
2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.		
3. <input checked="" type="checkbox"/> This is an express request to promptly begin national examination procedures (35 U.S.C. 371(f)).		
4. <input checked="" type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).		
5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2))		
a. <input checked="" type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau).		
b. <input checked="" type="checkbox"/> has been communicated by the International Bureau.		
c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).		
6. <input checked="" type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371 (c)(2)).		
7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))		
a. <input checked="" type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau).		
b. <input type="checkbox"/> have been communicated by the International Bureau.		
c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.		
d. <input type="checkbox"/> have not been made and will not be made.		
8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3))		
9. <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).		
"Executed"		
10. <input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (U.S.C. 371(c)(5)).		
<b>Items 11 to 16 below concern other document(s) or information included:</b>		
11. <input type="checkbox"/> Assignee: <u>MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD. of Osaka, JAPAN</u>		
12. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.		
13. <input checked="" type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.		
14. <input checked="" type="checkbox"/> A FIRST preliminary amendment.		
<input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.		
15. <input type="checkbox"/> A substitute specification.		
16. <input type="checkbox"/> A change of power of attorney and/or address letter.		
17. <input checked="" type="checkbox"/> Figure of Drawing to be published <u>1</u>		
18. <input checked="" type="checkbox"/> Other items or information:		
International Application as published in English and Japanese (Cover Sheets).		
Cover Letter Submitting Amended Pages of Application and Translation of Amended Pages (with five sheets in Japanese).		
Ten Sheets of Amended Pages.		
PCT/RO/101 (in Japanese).		
PCT/IB/304.		
PCT/IB/306.		
PCT/IB/308.		
PCT/IB/346.		
PCT/ISA/210.		
Cover Letter Regarding Executed Application.		
Claim of Priority.		

U.S. APPLICATION NO. (If known, see 37 CFR 1.5)

09/807427

INTERNATIONAL APPLICATION NO.

PCT/JP00/05621

ATTORNEY'S DOCKET NUMBER

P20934

19. The following fees are submitted:

Basic National Fee (37 CFR 1.492(a)(1)-(5)):

Search report has been prepared by the EPO or JPO. . . . . \$ 860.00

International preliminary examination fee paid to USPTO (37 CFR 1.482). . . . . \$ 690.00

No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO(37 CFR 1.445(a)(2)). . . . . \$ 710.00

Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO. . . . . \$1,000.00

International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4). . . . . \$ 100.00

ENTER APPROPRIATE BASIC FEE AMOUNT =

CALCULATIONS

PTO USE ONLY

\$860.00

Surcharge of \$130.00 for furnishing the oath or declaration later than 20 30 months from the earliest claimed priority date (37 CFR 1.492(e)).

\$

Claims

Number Filed

Number Extra

RATE

Total Claims

38

- 20 =

18

X \$18.00

\$324.00

Independent Claims

6

- 3 =

3

X \$80.00

\$240.00

Multiple dependent claim(s) (if applicable)

+ \$270.00

\$

TOTAL OF ABOVE CALCULATIONS =

\$1424.00

Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.

\$

SUBTOTAL =

\$1424.00

Processing fee of \$130.00 for furnishing the English translation later than 20 30 months from the earliest claimed priority date (37 CFR 1.492(f)).

+

Extension of Time fee in the amount of \$

TOTAL NATIONAL FEE =

\$1424.00

Fee for recording the enclosed assignment (37 CFR 1.21(h). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property

+

\$40.00

TOTAL FEES ENCLOSED =

\$1464.00

Amount to be refunded

\$

Charged

\$

a. ☒ A check in the amount of \$1464.00 to cover the above fees is enclosed.b. ☐ Please charge my Deposit Account No. \_\_\_\_\_ in the amount of \$\_\_\_\_\_ to cover the above fees.c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 19-0089.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO CUSTOMER NO. 7055

AT THE PRESENT ADDRESS OF:

Bruce H. Bernstein  
GREENBLUM & BERNSTEIN, P.L.C.  
1941 Roland Clarke Place  
Reston, VA 20191  
(703) 716-1191SIGNATURE  
Bruce H. Bernstein  
NAME29.027  
REGISTRATION NUMBER

P20934.A01

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Kazutoshi YASUNAGA et al.

Appl. No : Not Yet Assigned  
(National Stage of PCT/JP00/05621)

Filed : Concurrently Herewith (International Filing Date August 23, 2000)

For : APPARATUS AND METHOD FOR SPEECH CODING

**PRELIMINARY AMENDMENT AND COVER LETTER  
SUBMITTING AMENDED PAGES OF APPLICATION  
AND TRANSLATION OF AMENDED PAGES**Commissioner of Patents and Trademarks  
Washington, D.C. 20231

Sir:

Enclosed please find a copy of amendments to the claims under Article 19 which were received by the International Bureau on December 22, 2000, and an English translation of the amended claims.

Based upon the submission of the amended sheets of claims, Applicants respectfully request examination on the merits of the application containing amended pages of claims 1-38, in place of originally filed claims 1-14.

Additionally, prior to the examination of the above-identified application including replacement claims 1-38, amendment of claim 2 to correct typographical errors and claim 38 to remove multiple dependency, as follows, is respectfully requested.

IN THE CLAIMS

Please amend the claims as follows (a marked-up copy of the claim amendments is provided as an attachment to this Amendment):

2. (Amended) The speech encoder according to claim 1, wherein when one or more states of a previous subframe are an extremely large value or an extremely small value, said prediction coefficient adjusting means adjusts said prediction coefficients so as to reduce the influence thereof.

38. (Amended) A radio communication system that connects the communication base station according to claim 36 with a communication terminal via a radio network.


REMARKS

Entry of the foregoing amendment to the claims is respectfully requested prior to examination and calculation of the filing fees in the above-identified patent application.

Should there be any questions, the Examiner is invited to contact the undersigned at the below listed number.

Respectfully submitted,  
Kazutoshi YASUNAGA et al.

April 20, 2001  
GREENBLUM & BERNSTEIN, P.L.C.  
1941 Roland Clarke Place  
Reston, VA 20191  
(703) 716-1191

  
Bruce H. Bernstein  
Reg. No. 29,027  
33,329

MARKED-UP COPY OF AMENDED CLAIMS

2. (Amended) The speech encoder according to claim 1, wherein when one [ore] or more states of [aprevious subframes] a previous subframe are an extremely large value or an extremely small value, said prediction coefficient adjusting means adjusts said prediction coefficients so as to reduce the influence thereof.

38. (Amended) A radio communication system that connects the communication base station according to claim 36 with [the] a communication terminal [according to claim 37] via a radio network.

## DESCRIPTION

## APPARATUS AND METHOD FOR SPEECH CODING

## 5 Technical Field

The present invention relates to an apparatus and method for speech coding used in a digital communication system.

## 10 Background Art

In the field of digital mobile communication such as cellular telephones, there is a demand for a low bit rate speech compression coding method to cope with an increasing number of subscribers, and various research  
15 organizations are carrying forward research and development focused on this method.

In Japan, a coding method called "VSELP" with a bit rate of 11.2 kbps developed by Motorola, Inc. is used as a standard coding system for digital cellular  
20 telephones and digital cellular telephones using this system are on sale in Japan since the fall of 1994.

Furthermore, a coding system called "PSI-CELP" with a bit rate of 5.6 kbps developed by NTT Mobile Communications Network, Inc. is now commercialized.  
25 These systems are the improved versions of a system called "CELP" (described in "Code Excited Linear Prediction: M.R.Schroeder "High Quality Speech at Low Bit Rates", Proc.ICASSP '85, pp.937-940).

This CELP system is characterized by adopting a method (A-b-S: Analysis by Synthesis) consisting of separating speech into excitation information and vocal tract information, coding the excitation information using indices of a plurality of excitation samples stored in a codebook, while coding LPC (linear prediction coefficients) for the vocal tract information and making a comparison with input speech taking into consideration the vocal tract information during coding of the excitation information.

In this CELP system, an autocorrelation analysis and LPC analysis are conducted on the input speech data (input speech) to obtain LPC coefficients and the LPC coefficients obtained are coded to obtain an LPC code. The LPC code obtained is decoded to obtain decoded LPC coefficients. On the other hand, the input speech is assigned perceptual weight by a perceptual weighting filter using the LPC coefficients.

Two synthesized speeches are obtained by applying filtering to respective code vectors of excitation samples stored in an adaptive codebook and stochastic codebook (referred to as "adaptive code vector" (or adaptive excitation) and "stochastic code vector" (or stochastic excitation), respectively) using the obtained decoded LPC coefficients.

Then, a relationship between the two synthesized speeches obtained and the perceptual weighted input speech is analyzed, optimal values (optimal gains) of

the two synthesized speeches are obtained, the power of the synthesized speeches is adjusted according to the optimal gains obtained and an overall synthesized speech is obtained by adding up the respective synthesized  
 5 speeches. Then, coding distortion between the overall synthesized speech obtained and the input speech is calculated. In this way, coding distortion between the overall synthesized speech and input speech is calculated for all possible excitation samples and the  
 10 indexes of the excitation samples (adaptive excitation sample and stochastic excitation sample) corresponding to the minimum coding distortion are identified as the coded excitation samples.

The gains and indexes of the excitation samples  
 15 calculated in this way are coded and these coded gains and the indexes of the coded excitation samples are sent together with the LPC code to the transmission path. Furthermore, an actual excitation signal is created from two excitations corresponding to the gain code and  
 20 excitation sample index, these are stored in the adaptive codebook and at the same time the old excitation sample is discarded.

By the way, excitation searches for the adaptive codebook and for the stochastic codebook are generally  
 25 carried out on a subframe-basis, where subframe is a subdivision of an analysis frame. Coding of gains (gain quantization) is performed by vector quantization (VQ) that evaluates quantization distortion of the gains



using two synthesized speeches corresponding to the excitation sample indexes.

In this algorithm, a vector codebook is created beforehand which stores a plurality of typical samples (code vectors) of parameter vectors. Then, coding distortion between the perceptual weighted input speech and a perceptual weighted LPC synthesis of the adaptive excitation vector and of the stochastic excitation vector is calculated using gain code vectors stored in the vector codebook from the following expression 1:

$$En = \sum_{i=0}^I (Xi - gn \times Ai - hn \times Si)^2$$

Expression 1

where:

$E_n$  : Coding distortion when nth gain code vector is used

$X_i$  : Perceptual weighted speech

$A_i$  : Perceptual weighted LPC synthesis of adaptive code vector

$S_i$  : Perceptual weighted LPC synthesis of stochastic code vector

$g_n$  : Code vector element (gain on adaptive excitation side)

$h_n$  : Code vector element (gain on stochastic excitation side)

$n$  : Code vector number

$i$  : Excitation data index

$I$  : Subframe length (coding unit of input

speech) Then, distortion  $E_n$  when each code vector is used by controlling the vector codebook is compared and the number of the code vector with the least distortion is identified as the gain vector code. Furthermore, the number of the code vector with the least distortion is found from among all the possible code vectors stored in the vector codebook and identified to be the vector code.

Expression 1 above seems to require many computational complexity for every  $n$ , but since the sum of products on  $i$  can be calculated beforehand, it is possible to search  $n$  with a small amount of computational complexity.

On the other hand, by determining a code vector based on the transmitted code of the vector, a speech decoder (decoder) decodes coded data and obtains a code vector.

Moreover, further improvements have been made over the prior art based on the above algorithm. For example, taking advantage of the fact that the human perceptual characteristic to sound intensity is found to have logarithmic scale, power is logarithmically expressed and quantized, and two gains normalized with that power is subjected to VQ. This method is used in the Japan PDC half rate CODEC standard system. There is also a method of coding using inter-frame correlations of gain parameters (predictive coding). This method is used in the ITU-T international standard G.729. However, even

these improvements are unable to attain performance to a sufficient degree.

Gain information coding methods using the human perceptual characteristic to sound intensity and  
5 inter-frame correlations have been developed so far, providing more efficient coding performance of gain information. Especially, predictive quantization has drastically improved the performance, but the conventional method performs predictive quantization  
10 using the same values as those of previous subframes as state values. However, some of the values stored as state values are extremely large (small) and using those values for the next subframe may prevent the next subframe from being quantized correctly, resulting in  
15 local abnormal sounds.

#### Disclosure of Invention

It is an object of the present invention to provide a CELP type speech encoder and encoding method capable  
20 of performing speech encoding using predictive quantization with less including local abnormal sounds.

A subject of the present invention is to prevent local abnormal sounds by automatically adjusting prediction coefficients when the state value in a  
25 preceding subframe is an extremely large value or extremely small value in predictive quantization.

#### Brief Description of Drawings

FIG.1 is a block diagram showing a configuration of a radio communication apparatus equipped with a speech coder/decoder of the present invention;

FIG.2 is a block diagram showing a configuration of the speech encoder according to Embodiment 1 of the present invention;

FIG.3 is a block diagram showing a configuration of a gain calculation section of the speech encoder shown in FIG.2;

FIG.4 is a block diagram showing a configuration of a parameter coding section of the speech encoder shown in FIG.2;

FIG.5 is a block diagram showing a configuration of a speech decoder for decoding speech data coded by the speech encoder according to Embodiment 1 of the present invention;

FIG.6 is a drawing to explain an adaptive codebook search;

FIG.7 is a block diagram showing a configuration of a speech encoder according to Embodiment 2 of the present invention;

FIG.8 is a block diagram to explain a dispersed-pulse codebook;

FIG.9 is a block diagram showing an example of a detailed configuration of the dispersed-pulse codebook;

FIG.10 is a block diagram showing an example of a detailed configuration of the dispersed-pulse codebook;

FIG.11 is a block diagram showing a configuration of a speech encoder according to Embodiment 3 of the present invention;

FIG.12 is a block diagram showing a configuration of a speech decoder for decoding speech data coded by the speech coder according to Embodiment 3 of the present invention;

FIG.13A illustrates an example of a dispersed-pulse codebook used in the speech encoder according to Embodiment 3 of the present invention;

FIG.13B illustrates an example of the dispersed-pulse codebook used in the speech decoder according to Embodiment 3 of the present invention;

FIG.14A illustrates an example of the dispersed-pulse codebook used in the speech encoder according to Embodiment 3 of the present invention; and

FIG.14B illustrates an example of the dispersed-pulse codebook used in the speech decoder according to Embodiment 3 of the present invention.

20

Best Mode for Carrying out the Invention

With reference now to the attached drawings, embodiments of the present invention will be explained in detail below.

25 (Embodiment 1)

FIG.1 is a block diagram showing a configuration of a radio communication apparatus equipped with a speech encoder/decoder according to Embodiments 1 to 3 of the

present invention.

On the transmitting side of this radio communication apparatus, a speech is converted to an electric analog signal by speech input apparatus 11 such as a microphone and output to A/D converter 12. The analog speech signal is converted to a digital speech signal by A/D converter 12 and output to speech encoding section 13. Speech encoding section 13 performs speech encoding processing on the digital speech signal and outputs the coded information to modulation/demodulation section 14. Modulation/demodulation section 14 digital-modulates the coded speech signal and sends to radio transmission section 15. Radio transmission section 15 performs predetermined radio transmission processing on the modulated signal. This signal is transmitted via antenna 16. Processor 21 performs processing using data stored in RAM 22 and ROM 23 as appropriate.

On the other hand, on the receiving side of the radio communication apparatus, a reception signal received through antenna 16 is subjected to predetermined radio reception processing by radio reception section 17 and sent to modulation/demodulation section 14. Modulation/demodulation section 14 performs demodulation processing on the reception signal and outputs the demodulated signal to speech decoding section 18. Speech decoding section 18 performs decoding processing on the demodulated signal to obtain

a digital decoded speech signal and outputs the digital decoded speech signal to D/A converter 19. D/A converter 19 converts the digital decoded speech signal output from speech decoding section 18 to an analog decoded speech signal and outputs to speech output apparatus 20 such as a speaker. Finally, speech output apparatus 20 converts the electric analog decoded speech signal to a decoded speech and outputs the decoded speech.

Here, speech encoding section 13 and speech decoding section 18 are operated by processor 21 such as DSP using codebooks stored in RAM 22 and ROM 23. These operation programs are stored in ROM 23.

FIG.2 is a block diagram showing a configuration of a CELP type speech encoder according to Embodiment 1 of the present invention. This speech encoder is included in speech encoding section 13 shown in FIG.1. Adaptive codebook 103 shown in FIG.2 is stored in RAM 22 shown in FIG.1 and stochastic codebook 104 shown in FIG.2 is stored in ROM 23 shown in FIG.1.

In the speech encoder in FIG.2, LPC analysis section 102 performs an autocorrelation analysis and LPC analysis on speech data 101 and obtains LPC coefficients. Furthermore, LPC analysis section 102 performs encoding of the obtained LPC coefficients to obtain an LPC code. Furthermore, LPC analysis section 102 decodes the obtained LPC code and obtains decoded LPC coefficients. Speech data 101 input is sent to perceptual weighting section 107 and assigned perceptual weight using a

perceptual weighting filter using the LPC coefficients above.

Then, excitation vector generator 105 extracts an excitation vector sample (adaptive code vector or  
 5 adaptive excitation) stored in adaptive codebook 103 and an excitation vector sample (stochastic code vector or adaptive excitation) stored in stochastic codebook 104 and sends their respective code vectors to perceptual weighted LPC synthesis filter 106. Furthermore,  
 10 perceptual weighted LPC synthesis filter 106 performs filtering on the two excitation vectors obtained from excitation vector generator 105 using the decoded LPC coefficients obtained from LPC analysis section 102 and obtains two synthesized speeches.

15 Perceptual weighted LPC synthesis filter 106 uses a perceptual weighting filter using the LPC coefficients, high frequency enhancement filter and long-term prediction coefficient (obtained by carrying out a long-term prediction analysis of the input speech)  
 20 together and thereby performs a perceptual weighted LPC synthesis on their respective synthesized speeches.

Perceptual weighted LPC synthesis filter 106 outputs the two synthesized speeches to gain calculation section 108. Gain calculation section 108 has a  
 25 configuration shown in FIG.3. Gain calculation section 108 sends the two synthesized speeches obtained from perceptual weighted LPC synthesis filter 106 and the perceptual weighted input speech to analysis section



1081 and analyzes the relationship between the two synthesized speeches and input speech to obtain optimal values (optimal gains) for the two synthesized speeches. This optimal gains are output to power adjustment section  
 5 1082.

Power adjustment section 1082 adjusts the two synthesized speeches with the optimal gains obtained. The power-adjusted synthesized speeches are output to synthesis section 1083 and added up there to become an  
 10 overall synthesized speech. This overall synthesized speech is output to coding distortion calculation section 1084. Coding distortion calculation section 1084 finds coding distortion between the overall synthesized speech obtained and input speech.

15 Coding distortion calculation section 1084 controls excitation vector generator 105 to output all possible excitation vector samples of adaptive codebook 103 and of stochastic codebook 104, finds coding distortion between the overall synthesized speech and  
 20 input speech on all excitation vector samples and identifies the respective indexes of the respective excitation vector samples corresponding to the minimum coding distortion.

Then, analysis section 1081 sends the indexes of  
 25 the excitation vector samples, the two perceptual weighted LPC synthesized excitation vectors corresponding to the respective indexes and input speech to parameter coding section 109.

Parameter coding section 109 obtains a gain code by coding the gains and sends the LPC code, indexes of the excitation vector samples all together to the transmission path. Furthermore, parameter coding

5 section 109 creates an actual excitation vector signal from the gain code and two excitation vectors corresponding to the respective indexes and stores the excitation vector into the adaptive codebook 103 and at the same time discards the old excitation vector sample

10 in the adaptive codebook. By the way, an excitation vector search for the adaptive codebook and an excitation vector search for the stochastic codebook are generally performed on a subframe basis, where "subframe" is a subdivision of an processing frame(anàlysis frame).

15 Here, the operation of gain encoding of parameter coding section 109 of the speech encoder in the above configuration will be explained. FIG.4 is a block diagram showing a configuration of the parameter coding section of the speech encoder of the present invention.

20 In FIG.4, perceptual weighted input speech ( $X_i$ ), perceptual weighted LPC synthesized adaptive code vector ( $A_i$ ) and perceptual weighted LPC synthesized stochastic code vector ( $S_i$ ) are sent to parameter calculation section 1091. Parameter calculation section 1091

25 calculates parameters necessary for a coding distortion calculation. The parameters calculated by parameter calculation section 1091 are output to coding distortion calculation section 1092 and the coding distortion is

calculated there. This coding distortion is output to comparison section 1093. Comparison section 1093 controls coding distortion calculation section 1092 and vector codebook 1094 to obtain the most appropriate code  
 5 from the obtained coding distortion and outputs the code vector (decoded vector) obtained from vector codebook 1094 based on this code to decoded vector storage section 1096 and updates decoded vector storage section 1096.

Prediction coefficients storage section 1095  
 10 stores prediction coefficients used for predictive coding. This prediction coefficients are output to parameter calculation section 1091 and coding distortion calculation section 1092 to be used for parameter calculations and coding distortion calculations.  
 15 Decoded vector storage section 1096 stores the states for predictive coding. These states are output to parameter calculation section 1091 to be used for parameter calculations. Vector codebook 1094 stores code vectors.

20 Then, the algorithm of the gain coding method according to the present invention will be explained.

Vector codebook 1094 is created beforehand, which stores a plurality of typical samples (code vectors) of quantization target vectors. Each vector consists of  
 25 three elements; AC gain, logarithmic value of SC gain, and an adjustment coefficient for prediction coefficients of logarithmic value of SC gain.

This adjustment coefficient is a coefficient to

adjust prediction coefficients according to a states of previous subframes. More specifically, when a state of a previous subframe is an extremely large value or an extremely small value, this adjustment coefficient is set so as to reduce that influence. It is possible to calculate this adjustment coefficient using a training algorithm developed by the present inventor, et al. using many vector samples. Here, explanations of this training algorithm are omitted.

For example, a large value is set for the adjustment coefficient in a code vector frequently used for voiced sound segments. That is, when a same waveform is repeated in series, the reliability of the states of the previous subframes is high, and therefore a large adjustment coefficient is set so that the large prediction coefficients of the previous subframes can be used. This allows more efficient prediction.

On the other hand, a small value is set for the adjustment coefficient in a code vector less frequently used at the onset segments, etc. That is, when the waveform is quite different from the previous waveform, the reliability of the states of the previous subframes is low (the adaptive codebook is considered not to function), and therefore a small value is set for the adjustment coefficient so as to reduce the influence of the prediction coefficients of the previous subframes. This prevents any detrimental effect on the next prediction, making it possible to implement satisfactory

predictive coding.

In this way, adjusting prediction coefficients according to code vectors of states makes it possible to further improve the performance of predictive coding  
5 so far.

Prediction coefficients for predictive coding are stored in prediction coefficient storage section 1095. These prediction coefficients are prediction coefficients of MA (Moving Average) and two types of  
10 prediction coefficients, AC and SC, are stored by the number corresponding to the prediction order. These prediction coefficients are generally calculated through training based on a huge amount of sound database beforehand. Moreover, values indicating silent states  
15 are stored in decoded vector storage section 1096 as the initial values.

Then, the coding method will be explained in detail below. First, a perceptual weighted input speech ( $X_i$ ), perceptual weighted LPC synthesized adaptive code vector  
20 ( $A_i$ ) and perceptual weighted LPC synthesized stochastic code vector ( $S_i$ ) are sent to parameter calculation section 1091 and furthermore the decoded vector (AC, SC, adjustment coefficient) stored in decoded vector storage section 1096 and the prediction coefficients (AC, SC)  
25 stored in prediction coefficient storage section 1095 are sent. Parameters necessary for a coding distortion calculation are calculated using these values and vectors.

A coding distortion calculation by coding distortion calculation section 1092 is performed according to expression 2 below:

$$E_n = \sum_{i=0}^I (X_i - G_{an} \times A_i - G_{sn} \times S_i)^2$$

5

Expression 2

where:

$G_{an}$ ,  $G_{sn}$ : Decoded gain

$E_n$ : Coding distortion when nth gain code vector is

10 used

$X_i$ : Perceptual weighted speech

$A_i$ : Perceptual weighted LPC synthesized adaptive code vector

$S_i$ : Perceptual weighted LPC synthesized stochastic  
15 code vector

$n$ : Code vector number

$i$ : Excitation vector index

$I$ : Subframe length (coding unit of input speech)

In order to reduce the amount of calculation,  
20 parameter calculation section 1091 calculates the part independent of the code vector number. What should be calculated are correlations between three synthesized speeches ( $X_i$ ,  $A_i$ ,  $S_i$ ) and powers. These calculations are performed according to expression 3 below:

$$D_{xx} = \sum_{i=0}^I X_i \times X_i$$

$$D_{xa} = \sum_{i=0}^I X_i \times A_i \times 2$$

$$D_{xs} = \sum_{i=0}^I X_i \times S_i \times 2$$

$$D_{aa} = \sum_{i=0}^I A_i \times A_i$$

$$D_{as} = \sum_{i=0}^I A_i \times S_i \times 2$$

$$D_{ss} = \sum_{i=0}^I S_i \times S_i$$

### Expression 3

where:

5  $D_{xx}, D_{xa}, D_{xs}, D_{aa}, D_{as}, D_{ss}$  : Correlation value between  
synthesized speeches, power

$X_i$  : Perceptual weighted speech

$A_i$  : Perceptual weighted LPC synthesized adaptive  
code vector

10  $S_i$  : Perceptual weighted LPC synthesized stochastic  
code vector

$n$  : Code vector number

$i$  : Excitation vector index

$I$  : Subframe length (coding unit of input speech)

Furthermore, parameter calculation section 1091  
15 calculates three predictive values shown in expression  
4 below using past code vectors stored in decoded vector  
storage section 1096 and prediction coefficients stored  
in prediction coefficient storage section 1095.

$$Pr a = \sum_{m=0}^M \alpha_m \times S_{am}$$

$$Pr s = \sum_{m=0}^M \beta_m \times S_{cm} \times S_{sm}$$

$$P_{sc} = \sum_{m=0}^M \beta_m \times S_{cm}$$

Expression 4

where:

$P_{ra}$ : Predictive value (AC gain)

5  $P_{rs}$ : Predictive value (SC gain)

$P_{sc}$ : Predictive value (prediction coefficient)

$\alpha_m$ : Prediction coefficient (AC gain, fixed value)

$\beta_m$ : Prediction coefficient (SC gain, fixed value)

$S_{am}$ : State (element of past code vector, AC gain)

10  $S_{sm}$ : State (element of past code vector, SC gain)

$S_{cm}$ : State (element of past code vector, SC

prediction coefficient adjustment coefficient)

$m$ : Predictive index

$M$ : Prediction order

15 As is apparent from expression 4 above, with regard to  $P_{rs}$  and  $P_{sc}$ , adjustment coefficients are multiplied unlike the conventional art. Therefore, regarding the predictive value and prediction coefficient of an SC gain, when a value of a state in the previous subframe is

20 extremely large or extremely small, it is possible to alleviate the influence (reduce the influence) by means of the adjustment coefficient. That is, it is possible to adaptively change the predictive value and prediction coefficients of the SC gain according to the states.



Then, coding distortion calculation section 1092 calculates coding distortion using the parameters calculated by parameter calculation section 1091, the prediction coefficients stored in prediction

5 coefficient storage section 1095 and the code vectors stored in vector codebook 1094 according to expression 5 below:

$$En = D_{xx} + (G_{an})^2 \times D_{aa} + (G_{sn})^2 \times D_{ss} - G_{an} \times D_{xa} - G_{sn} \times D_{xs} + G_{an} \times G_{sn} \times D_{as}$$

$$G_{an} = Pr_a + (1 - Pac) \times Can$$

$$G_{sn} = 10^{\{Pr_s + (1 - Psc) \times Csn\}}$$

Expression 5

10 where:

$E_n$ : Coding distortion when nth gain code vector is used

$D_{xx}, D_{xa}, D_{xs}, D_{aa}, D_{as}, D_{ss}$ : Correlation value between synthesized speeches, power

15  $G_{an}, G_{sn}$ : Decoded gain

$Pr_a$ : Predictive value (AC gain)

$Pr_s$ : Predictive value (SC gain)

$Pac$ : Sum of prediction coefficients (fixed value)

20  $Psc$ : Sum of prediction coefficients (calculated by expression 4 above)

$C_{an}, C_{sn}, C_{cn}$ : Code vector,  $C_{cn}$  is a prediction coefficient adjustment coefficient, but not used here

$n$ : Code vector number

$D_{xx}$  is actually independent of code vector number  $n$ , and the addition of  $D_{xx}$  can be omitted.

Then, comparison section 1093 controls vector codebook 1094 and coding distortion calculation section 1092 and finds the code vector number corresponding to the minimum coding distortion calculated by coding distortion calculation section 1092 from among a plurality of code vectors stored in vector codebook 1094 and identifies this as the gain code. Furthermore, the content of decoded vector storage section 1096 is updated using the gain code obtained. The update is performed according to expression 6 below:

$$\begin{aligned} S_{am} &= S_{am} - 1(m=M \sim 1), S_{a0} = C_{aJ} \\ S_{sm} &= S_{sm} - 1(m=M \sim 1), S_{s0} = C_{sJ} \\ S_{cm} &= S_{cm} - 1(m=M \sim 1), S_{c0} = C_{cJ} \end{aligned}$$

Expression 6

where:

$S_{am}$ ,  $S_{sm}$ ,  $S_{cm}$ : State vector (AC, SC, prediction coefficient adjustment coefficient)

$m$ : Predictive index

$M$ : Prediction order

$J$ : Code obtained from comparison section

As is apparent from Expression 4 to Expression 6, in this embodiment, decoded vector storage section 1096 stores state vector  $S_{cm}$  and prediction coefficients are adaptively controlled using these prediction coefficient adjustment coefficients.

FIG.5 shows a block diagram showing a configuration of the speech decoder according to this embodiment of the present invention. This speech decoder is included

in speech decoding section 18 shown in FIG.1. By the way, adaptive codebook 202 in FIG.5 is stored in RAM 22 in FIG.1 and stochastic codebook 203 in FIG.5 is stored in ROM 23 in FIG.1.

- 5           In the speech decoder in FIG.5, parameter decoding section 201 obtains the respective excitation vector sample codes of respective excitation vector codebooks (adaptive codebook 202, stochastic codebook 203), LPC codes and gain codes from the transmission path.
- 10          Parameter decoding section 201 then obtains decoded LPC coefficients from the LPC code and obtains decoded gains from the gain code.

- Then, excitation vector generator 204 obtains decoded excitation vectors by multiplying the
- 15          respective excitation vector samples by the decoded gains and adding up the multiplication results. In this case, the decoded excitation vector obtained are stored in adaptive codebook 204 as excitation vector samples and at the same time the old excitation vector samples
- 20          are discarded. Then, LPC synthesis section 205 obtains a synthesized speech by filtering the decoded excitation vector with the decoded LPC coefficients.

- The two excitation codebooks are the same as those included in the speech encoder in FIG.2 (reference
- 25          numerals 103 and 104 in FIG.2) and the sample numbers (codes for the adaptive codebook and codes for the stochastic codebook) to extract the excitation vector samples are supplied from parameter decoding section

201.

Thus, the speech encoder of this embodiment can control prediction coefficients according to each code vector, providing more efficient prediction more  
 5 adaptable to local characteristic of speech, thus making it possible to prevent detrimental effects on prediction in the non-stationary segment and attain special effects that have not been attained by conventional arts.

(Embodiment 2)

10 As described above, the gain calculation section in the speech encoder compares synthesized speeches and input speeches of all possible excitation vectors in the adaptive codebook and in the stochastic codebook obtained from the excitation vector generator. At this  
 15 time, two excitation vectors (adaptive codebook vector and stochastic codebook vector) are generally searched in an open-loop for the consideration of the amount of computational complexity. This will be explained with reference to FIG.2 below.

20 In this open-loop search, excitation vector generator 105 selects excitation vector candidates only from adaptive codebook 103 one after another, makes perceptual weighted LPC synthesis filter 106 function to obtain a synthesized speech and send to gain  
 25 calculation section 108, compares the synthesized speech and input speech and selects an optimal code of adaptive codebook 103.

Then, excitation vector generator 105 fixes the

code of adaptive codebook 103 above, selects the same excitation vector from adaptive codebook 103 and selects excitation vectors corresponding to gain calculation section 108 one after another from stochastic codebook 104 and sends to perceptual weighted LPC synthesis filter 106. Gain calculation section 108 compares the sum of both synthesized speeches and the input speech to determine the code of stochastic codebook 104.

When this algorithm is used, the coding performance deteriorates slightly compared to searching codes of all codebooks respectively, but the amount of computational complexity is reduced drastically. For this reason, this open-loop search is generally used.

Here, a typical algorithm in a conventional open-loop excitation vector search will be explained. Here, the excitation vector search procedure when one analysis section (frame) is composed of two subframes will be explained.

First, upon reception of an instruction from gain calculation section 108, excitation vector generator 105 extracts an excitation vector from adaptive codebook 103 and sends to perceptual weighted LPC synthesis filter 106. Gain calculation section 108 repeatedly compares the synthesized excitation vector and the input speech of the first subframe to find an optimal code. Here, the features of the adaptive codebook will be shown. The adaptive codebook consists of excitation vectors past used for speech synthesis. A code corresponds to a time

lag as shown in FIG.6.

Then, after a code of adaptive codebook 103 is determined, a search for the stochastic codebook is started. Excitation vector generator 105 extracts the  
 5 excitation vector of the code obtained from the search of the adaptive codebook 103 and the excitation vector of the stochastic codebook 104 specified by gain calculation section 108 and sends these excitation vectors to perceptual weighted LPC synthesis filter 106.  
 10 Then, gain calculation section 108 calculates coding distortion between the perceptual weighted synthesis speech and perceptual weighted input speech and determines an optimal (whose square error becomes a minimum) code of stochastic excitation vector 104. The  
 15 procedure for an excitation vector code search in one analysis section (in the case of two subframes) is shown below.

- 1) Determines the code of the adaptive codebook of the first subframe.
- 20 2) Determines the code of the stochastic codebook of the first subframe.
- 3) Parameter coding section 109 codes gains, generates the excitation vector of the first subframe with decoded gains and updates adaptive codebook 103.
- 25 4) Determines the code of the adaptive codebook of the second subframe.
- 5) Determines the code of the stochastic codebook of the second subframe.

6) Parameter coding section 109 codes the gains, generates the excitation vector of the second subframe with decoded gain and updates adaptive codebook 103.

The algorithm above allows efficient coding of  
 5 excitation vectors. However, an effort has been recently developed for decreasing the number of bits of excitation vectors aiming at a further reduction of the bit rate. What receives special attention is an algorithm of reducing the number of bits by taking  
 10 advantage of the presence of a large correlation in a lag of the adaptive codebook and narrowing the search range of the second subframe to the range close to the lag of the first subframe (reducing the number of entries) while leaving the code of the first subframe  
 15 as it is.

With this recently developed algorithm, local deterioration may be provoked, in the case speech signal in an analysis segment (frame) has a large change, or in the case the characteristics of the consecutive two  
 20 frames are much different

This embodiment provides a speech encoder that implements a search method of calculating correlation values by performing a pitch analysis for two subframes respectively, before starting coding and determining the  
 25 range of searching a lag between two subframes based on the correlation values obtained.

More specifically, the speech encoder of this embodiment is a CELP type encoder that breaks down one

frame into a plurality of subframes and codes respective frames, characterized by comprising a pitch analysis section that performs a pitch analysis of a plurality of subframes in the processing frame respectively, and

5 calculates correlation values before searching the first subframe in the adaptive codebook and a search range setting section that while the pitch analysis section calculates correlation values of a plurality of

10 subframes in the processing frame respectively, finds the value most likely to be the pitch cycle (typical pitch) on each subframe from the size of the correlation values and determines the search range of a lag between a plurality of subframes based on the correlation values obtained by the pitch analysis section and the typical

15 pitch. Then, the search range setting section of this speech encoder determines a provisional pitch that becomes the center of the search range using the typical pitch of a plurality of subframes obtained by the pitch analysis section and the correlation value and the search

20 range setting section sets the lag search range in a specified range around the determined provisional pitch and sets the search range before and after the provisional pitch when the lag search range is set. Moreover, in this case, the search range setting section

25 reduces the number of candidates for the short lag section (pitch period), widely sets the range of a long lag and searches the lag in the range set by the search range setting section during the search in the adaptive



codebook.

The speech encoder of this embodiment will be explained in detail below using the attached drawings. Here, suppose one frame is divided into two subframes.

- 5 The same procedure can also be used for coding in the case of 3 subframes or more.

In a pitch search according to a so-called delta lag coding system, this speech coder finds pitches of all subframes in the processing frame, determines the level of a correlation between pitches and determines the search range according to the correlation result.

FIG.7 is a block diagram showing a configuration of the speech encoder according to Embodiment 2 of the present invention. First, LPC analysis section 302 performs an autocorrelation analysis and LPC analysis on speech data input (input speech) 301 entered and obtains LPC coefficients. Moreover, LPC analysis section 302 performs coding on the LPC coefficients obtained and obtains an LPC code. Furthermore, LPC analysis section 302 decodes the LPC code obtained and obtains decoded LPC coefficients.

Then, pitch analysis section 310 performs pitch analysis for consecutive 2 subframe respectively, and obtains a pitch candidate and a parameter for each subframe. The pitch analysis algorithm for one subframe is shown below. Two correlation coefficients are obtained from expression 7 below. At this time,  $C_{pp}$  is obtained about  $P_{min}$  first and remaining  $P_{min+1}$  and  $P_{min+2}$  can

be calculated efficiently by subtraction and addition of the values at the frame end.

$$V_p = \sum_{i=0}^L X_i \times X_i - P \quad (P=P_{\min} \sim P_{\max})$$

$$C_{pp} = \sum_{i=0}^L X_i - P \times X_i - P \quad (P=P_{\min} \sim P_{\max})$$

5

Expression 7

where:

$XX_i, X_{i-P}$ : Input speech

$V_p$ : Autocorrelation function

$C_{pp}$ : Power component

10 

$i$ : Input speech sample number

$L$ : Subframe length

$P$ : Pitch

$P_{\min}, P_{\max}$ : Minimum value and maximum value for pitch search

15

Then, the autocorrelation function and power component calculated from expression 7 above are stored in memory and the following procedure is used to calculate typical pitch  $P_1$ . This is the processing of calculating pitch  $P$  that corresponds to a maximum of  $V_p \times V_p / C_{pp}$  while  $V_p$  is positive. However, since a division calculation generally requires a greater amount of computational complexities, both the numerator and denominator are stored to convert the division to a multiplication to reduce the computational complexities.

25

Here, a pitch is found in such a way that the sum of square of the input speech and the square of the

difference between the input speech and the adaptive excitation vector ahead of the input speech by the pitch becomes a minimum. This processing is equivalent to the processing of finding pitch  $P$  corresponding to a maximum

5 of  $V_p \times V_p / C_{pp}$ . Specific processing is as follows:

- 1) Initialization ( $P=P_{\min}$ ,  $VV=C=0$ ,  $P_1=P_{\min}$ )
- 2) If ( $V_p \times V_p \times C < VV \times C_{pp}$ ) or ( $V_p < 0$ ), then go to 4). Otherwise, go to 3).
- 3) Supposing  $VV=V_p \times V_p$ ,  $C=C_{pp}$ ,  $P_1=P$ , go to 4).
- 10 4) Suppose  $P=P+1$ . At this time, if  $P > P_{\max}$ , the process ends. Otherwise, go to 2).

Perform the operation above for each of 2 subframes to calculate typical pitches  $P_1$  and  $P_2$ , autocorrelation coefficients  $V_{1p}$  and  $V_{2p}$ , power components  $C_{1pp}$  and  $C_{2pp}$

15 ( $P_{\min} < p < P_{\max}$ ).

Then, search range setting section 311 sets the search range of the lag in the adaptive codebook. First, a provisional pitch, which is the center of the search range is calculated. The provisional pitch is

20 calculated using the typical pitch and parameter obtained by pitch analysis section 310.

Provisional pitches  $Q_1$  and  $Q_2$  are calculated using the following procedure. In the following explanation, constant  $Th$  (more specifically, a value 6 or so is

25 appropriate) as the lag range. Moreover, the correlation value obtained from expression 7 above is used.

While  $P_1$  is fixed, provisional pitch ( $Q_2$ ) with the

maximum correlation is found near  $P_1$  ( $\pm Th$ ) first.

1) Initialization ( $p=P_1-Th$ ,  $C_{max}=0$ ,  $Q_1=P_1$ ,  $Q_2=P_1$ )

2) If  $(V_{1p1} \times V_{1p1} / C_{1p1p1} + V_{2p} \times V_{2p} / C_{2pp} < C_{max})$  or  $(V_{2p} < 0)$  then go to 4). Otherwise, go to 3).

5 3) Supposing  $C_{max}=V_{1p1} \times V_{1p1} / C_{1p1p1} + V_{2p} \times V_{2p} / C_{2pp}$ ,  $Q_2=p$ , go to 4).

4) Supposing  $p=p+1$ , go to 2). However, at this time, if  $p > P_1 + Th$ , go to 5).

In this way, processing in 2) to 4) is performed  
10 from  $P_1 - Th$  to  $P_1 + Th$ , the one with the maximum correlation,  $C_{max}$  and provisional pitch  $Q_2$  are found.

Then, while  $P_2$  is fixed, provisional pitch ( $Q_1$ ) near  $P_2$  ( $\pm Th$ ) with a maximum correlation is found. In this case,  $C_{max}$  will not be initialized. By calculating  $Q_1$   
15 whose correlation becomes a maximum including  $C_{max}$  when  $Q_2$  is found, it is possible to find  $Q_1$  and  $Q_2$  with the maximum correlation between the first and second subframes.

5) Initialization ( $p=P_2-Th$ )

20 6) If  $(V_{1p} \times V_{1p} / C_{1pp} + V_{2p2} \times V_{2p2} / C_{2p2p2} < C_{max})$  or  $(V_{1p} < 0)$ , go to 8). Otherwise, go to 7).

7) Supposing

$C_{max}=V_{1p} \times V_{1p} / C_{1pp} + V_{2p2} \times V_{2p2} / C_{2p2p2}$ ,  $Q_1=p$ ,  $Q_2=P_2$ , go to 8).

8) Supposing  $p=p+1$ , go to 6). However, at this time  
25 if  $p > P_2 + Th$ , go to 9).

9) End

In this way, perform processing in 6) to 8) from  $P_2 - Th$  to  $P_2 + Th$ , the one with the maximum correlation,  $C_{max}$

and provisional pitches  $Q_1$  and  $Q_2$  are found.  $Q_1$  and  $Q_2$  at this time are provisional pitches of the first and second subframes, respectively.

From the algorithm above, it is possible to select  
 5 two provisional pitches with a relatively small difference in size (the maximum difference is  $Th$ ) while evaluating the correlation between two subframes simultaneously. Using these provisional pitches prevents the coding performance from drastically  
 10 deteriorating even if a small search range is set during a search of the second subframe in the adaptive codebook. For example, when sound quality changes suddenly from the second subframe, if there is a strong correlation of the second subframe, using  $Q_1$  that reflects the  
 15 correlation of the second subframe can avoid the deterioration of the second subframe.

Furthermore, search range setting section 311 sets the search range ( $L_{ST}$  to  $L_{EN}$ ) of the adaptive codebook using provisional pitch  $Q_1$  obtained as expression 8  
 20 below:

First subframe

$$L_{ST}=Q_1-5 \quad (\text{when } L_{ST}<L_{min}, L_{ST}=L_{min})$$

$$L_{EN}=L_{ST}+20 \quad (\text{when } L_{ST}>L_{max}, L_{ST}=L_{max})$$

Second subframe

$$25 \quad L_{ST}=T_1-10 \quad (\text{when } L_{ST}<L_{min}, L_{ST}=L_{min})$$

$$L_{EN}=L_{ST}+21 \quad (\text{when } L_{ST}>L_{max}, L_{ST}=L_{max})$$

Expression 8

where:

$L_{ST}$ : Minimum of search range

$L_{EN}$ : Maximum of search range

$L_{min}$ : Minimum value of lag (e.g., 20)

$L_{max}$ : Maximum value of lag (e.g., 143)

5       $T_1$ : Adaptive codebook lag of first frame

In the above setting, it is not necessary to narrow the search range for the first subframe. However, the present inventor, et al. have confirmed through experiments that the performance is improved by setting the vicinity of a value based on the pitch of the input speech as the search range and this embodiment uses an algorithm of searching by narrowing the search range to 26 samples.

On the other hand, for the second subframe, the search range is set to the vicinity of lag  $T_1$  obtained by the first subframe. Therefore, it is possible to perform 5-bit coding on the adaptive codebook lag of the second subframe with a total of 32 entries. Furthermore, the present inventor, et al. have also confirmed this time through experiments that the performance is improved by setting fewer candidates with a short lag and more candidates with a long lag. However, as is apparent from the explanations heretofore, this embodiment does not use provisional pitch  $Q_2$ .

25      Here, the effects of this embodiment will be explained. In the vicinity of the provisional pitch of the first subframe obtained by search range setting section 311, the provisional pitch of the second subframe

also exists (because it is restricted with constant  $T_h$ ). Furthermore, since a search has been performed with the search range narrowed in the first subframe, the lag resultant from the search is not separated from the  
 5 provisional pitch of the first subframe.

Therefore, when the second subframe is searched, the search can be performed in the range close to the provisional pitch of the second subframe, and therefore it is possible to search lags appropriate for both the  
 10 first and second frames.

Suppose a example where the first subframe is a silent-speech and the second subframe is not a silent-speech. According to the conventional method, sound quality will deteriorate drastically if the second  
 15 subframe pitch is no longer included in the search section by narrowing the search range. According to the method of this embodiment, a strong correlation of typical pitch  $P_2$  is reflected in the analysis of the provisional pitch of the pitch analysis section.

Therefore, the provisional pitch of the first subframe has a value close to  $P_2$ . This makes it possible to determine the range close to the part at which the speech starts as the provisional pitch in the case of a search by a delta lag. That is, in the case of an adaptive  
 20 codebook search of the second subframe, a value close to  $P_2$  can be searched, and therefore it is possible to perform an adaptive codebook search of the second subframe by a delta lag even if speech starts at some

midpoint in the second subframe.

Then, excitation vector generator 305 extracts the excitation vector sample (adaptive code vector or adaptive excitation vector) stored in adaptive codebook 303 and the excitation vector sample (stochastic code vector or stochastic excitation vector) stored in stochastic codebook 304 and sends these excitation vector samples to perceptual weighted LPC synthesis filter 306. Furthermore, perceptual weighted LPC synthesis filter 306 performs filtering on the two excitation vectors obtained by excitation vector generator 305 using the decoded LPC coefficients obtained by LPC analysis section 302.

Furthermore, gain calculation section 308 analyzes the relationship between the two synthesized speeches obtained by perceptual weighted LPC synthesis filter 306 and the input speech and finds respective optimal values (optimal gains) of the two synthesized speeches. Gain calculation section 308 adds up the respective synthesized speeches with power adjusted with the optimal gain and obtains an overall synthesized speech. Then, gain calculation section 308 calculates coding distortion between the overall synthesized speech and the input speech. Furthermore, gain calculation section 308 calculates coding distortion between many synthesized speeches obtained by making function excitation vector generator 305 and perceptual weighted LPC synthesis filter 306 on all excitation vector samples



in adaptive codebook 303 and stochastic codebook 304 and the input speech, and finds the indexes of the excitation vector samples corresponding to the minimum of the resultant coding distortion.

5           Then, gain calculation section 308 sends the indexes of the excitation vector samples obtained and the two excitation vectors corresponding to the indexes and the input speech to parameter coding section 309. Parameter coding section 309 obtains a gain code by  
10 performing gain coding and sends the gain code together with the LPC code and indexes of the excitation vector samples to the transmission path.

          Furthermore, parameter coding section 309 creates an actual excitation vector signal from the gain code  
15 and the two excitation vectors corresponding to the indexes of the excitation vector samples and stores the actual excitation vector signal in adaptive codebook 303 and at the same time discards the old excitation vector sample.

20           By the way, perceptual weighted LPC synthesis filter 306 uses a perceptual weighting filter using an LPC coefficients, high frequency enhancement filter and long-term prediction coefficient (obtained by performing a long-term predictive analysis of the input  
25 speech).

          Gain calculation section 308 above makes a comparison with the input speech about all possible excitation vectors in adaptive codebook 303 and all

possible stochastic codebook 304 obtained from  
 excitation vector generator 305, but two excitation  
 vectors (adaptive codebook 303 and stochastic codebook  
 304) are searched in an openloop as described above in  
 5 order to reduce the amount of computational complexity.

Thus, the pitch search method in this embodiment  
 performs pitch analyses of a plurality of subframes in  
 the processing frame respectively before performing an  
 adaptive codebook search of the first subframe, then  
 10 calculates a correlation value and thereby can control  
 correlation values of all subframes in the frame  
 simultaneously.

Then, the pitch search method in this embodiment  
 calculates a correlation value of each subframe, finds  
 15 a value most likely to be a pitch period (called a "typical  
 pitch") in each subframe according to the size of the  
 correlation value and sets the lag search range of a  
 plurality of subframes based on the correlation value  
 obtained from the pitch analysis and typical pitch. In  
 20 the setting of this search range, the pitch search method  
 in this embodiment obtains an appropriate provisional  
 pitch (called a "provisional pitch") with a small  
 difference, which will be the center of the search range,  
 using the typical pitches of a plurality of subframes  
 25 obtained from the pitch analyses and the correlation  
 values.

Furthermore, the pitch search method in this  
 embodiment confines the lag search section to a specified

range before and after the provisional pitch obtained in the setting of the search range above, allowing an efficient search of the adaptive codebook. In that case, the pitch search method in this embodiment sets fewer  
 5 candidates with a short lag part and a wider range with a long lag, making it possible to set an appropriate search range where satisfactory performance can be obtained. Furthermore, the pitch search method in this embodiment performs a lag search within the range set  
 10 by the setting of the search range above during an adaptive codebook search, allowing coding capable of obtaining satisfactory decoded sound.

Thus, according to this embodiment, the provisional pitch of the second subframe also exists near  
 15 the provisional pitch of the first subframe obtained by search range setting section 311 and the search range is narrowed in the first subframe, and therefore the lag resulting from the search does not get away from the provisional pitch. Therefore, during a search of the  
 20 second subframe, it is possible to search around the provisional pitch of the second subframe allowing an appropriate lag search in the first and second subframes even in a non-stationary frame in the case where a speech starts from the last half of a frame, and thereby attain  
 25 a special effect that has not been attained with conventional arts.

(Embodiment 3)

An initial CELP system uses a stochastic codebook

with entries of a plurality of types of random sequence as stochastic excitation vectors, that is, a stochastic codebook with a plurality of types of random sequence directly stored in memory. On the other hand, many low  
 5 bit-rate CELP encoder/decoder have been developed in recent years, which include an algebraic codebook to generate stochastic excitation vectors containing a small number of non-zero elements whose amplitude is +1 or -1 (the amplitude of elements other than the non-  
 10 zero element is zero) in the stochastic codebook section.

By the way, the algebraic codebook is disclosed in the "Fast CELP Coding based on Algebraic codes", J.Adoul et al, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 1987, pp. 1957-1960 or "Comparison of Some  
 15 Algebraic Structure for CELP Coding of Speech", J.Adoul et al, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 1987, pp. 1953-1956, etc.

The algebraic codebook disclosed in the above papers is a codebook having excellent features such as  
 20 (1) ability to generate synthesized speech of high quality when applied to a CELP system with a bit rate of approximately 8 kb/s, (2) ability to search a stochastic with a small amount of computational complexity, and (3) elimination of the necessity of data  
 25 ROM capacity to directly store stochastic excitation vectors.

Then, CS-ACELP (bit rate: 8 kb/s) and ACELP (bit rate: 5.3 kb/s) characterized by using an algebraic

codebook as a stochastic codebook are recommended as G.729 and g723.1, respectively from the ITU-T in 1996. By the way, detailed technologies of CS-ACELP are disclosed in "Design and Description of CS-ACELP: A Toll  
 5 Quality 8 kb/s Speech Coder", Redwan Salami et al, IEEE trans. SPEECH AND AUDIO PROCESSING, vol. 6, no. 2, March 1998, etc.

The algebraic codebook is a codebook with the excellent features as described above. However, when  
 10 the algebraic codebook is applied to the stochastic codebook of a CELP encoder/decoder, the target vector for stochastic codebook search is always encoded/decoded (vector quantization) with stochastic excitation vectors including a small number of non-zero elements,  
 15 and thus the algebraic codebook has a problem that it is impossible to express a target vector for stochastic codebook search in high fidelity. This problem becomes especially conspicuous when the processing frame corresponds to an unvoiced consonant  
 20 segment or background noise segment.

This is because the target vector for stochastic codebook search often takes a complicated shape in an unvoiced consonant segment or background noise segment. Furthermore, in the case where the algebraic codebook  
 25 is applied to a CELP encoder/decoder whose bit rate is much lower than the order of 8 kb/s, the number of non-zero elements in the stochastic excitation vector is reduced, and therefore the above problem can become a bottleneck

even in a stationary voiced segment where the target vector for stochastic codebook search is likely to be a pulse-like shape.

As one of methods for solving the above problem of the algebraic codebook, a method using a dispersed-pulse codebook is disclosed, which uses a vector obtained by convoluting a vector containing a small number of non-zero elements (elements other than non-zero elements have a zero value) output from the algebraic codebook and a fixed waveform called a "dispersion pattern" as the excitation vector of a synthesis filter. The dispersed-pulse codebook is disclosed in the Unexamined Japanese Patent Publication No.HEI 10-232696, "ACELP Coding with Dispersed-Pulse Codebook" (by Yasunaga, et al., Collection of Preliminary Manuscripts of National Conference of Institute of Electronics, Information and Communication Engineers in Springtime 1997, D-14-11, p.253, 1997-03) and "A Low Bit Rate Speech Coding with Multi Dispersed Pulse based Codebook" (by Yasunaga, et al., Collected Papers of Research Lecture Conference of Acoustical Society of Japan in Autumn 1998, pp.281-282, 1998-10), etc.

Next, an outline of the dispersed-pulse codebook disclosed in the above papers will be explained using FIG.8 and FIG.9. FIG.9 shows a further detailed example of the dispersed-pulse codebook in FIG.8.

In the dispersed-pulse codebook in FIG.8 and FIG.9, algebraic codebook 4011 is a codebook for generating a

pulse vector made up of a small number of non-zero elements (amplitude is +1 or -1). The CELP encoder/decoder described in the above papers uses a pulse vector (made up of a small number of non-zero  
 5 elements), which is the output of algebraic codebook 4011, as the stochastic excitation vector.

Dispersion pattern storage section 4012 stores at least one type of fixed waveform called a "dispersion pattern" for every channel. There can be two cases of  
 10 dispersion patterns stored for every channel: one case where dispersion patterns differing from one channel to another are stored and the other case where a dispersion pattern of a same (common) shape for all channels is stored. The case where a common dispersion pattern is  
 15 stored for all channels corresponds to simplification of the case where dispersion pattern differing from one channel to another are stored, and therefore the case where dispersion patterns differing from one channel to another are stored will be explained in the following  
 20 explanations of the present description.

Instead of directly outputting the output vector from algebraic codebook 4011 as a stochastic excitation vector, dispersed-pulse codebook 401 convolutes the vector output from algebraic codebook 4011 and  
 25 dispersion patterns read from dispersion pattern storage section 4012 for every channel in pulse dispersing section 4013, adds up vectors resulting from the convolution calculations and uses the resulting vector

as the stochastic excitation vector.

The CELP encoder/decoder disclosed in the above papers is characterized by using a dispersed-pulse codebook in a same configuration for the encoder and  
 5 decoder (the number of channels in the algebraic codebook, the number of types and shape of dispersion patterns registered in the dispersion pattern storage section are common between the encoder and decoder). Moreover, the CELP encoder/decoder disclosed in the above papers aims  
 10 at improving the quality of synthesized speech by efficiently setting the shapes and the number of types of dispersion patterns registered in dispersion pattern storage section 4012, and the method of selecting in the case where a plurality of types of dispersion patterns  
 15 are registered.

By the way, the explanation of the dispersed-pulse codebook here describes the case where an algebraic codebook that confines the amplitude of non-zero elements to +1 or -1 is used as the codebook for generating  
 20 a pulse vector made up of a small number of non-zero elements. However, as the codebook for generating the relevant pulse vectors, it is also possible to use a multi-pulse codebook that does not confine the amplitude of non-zero elements or a regular pulse codebook, and  
 25 in such cases, it is also possible to improve the quality of the synthesized speech by using a pulse vector convoluted with a dispersion pattern as the stochastic excitation vector.



It has been disclosed so far that it is possible to effectively improve the quality of a synthesized speech by registering dispersion patterns obtained by statistically training of shapes based on a huge number of target vectors for stochastic codebook search, dispersion patterns of random-like shapes to efficiently express the unvoiced consonant segments and noise-like segments, dispersion patterns of pulse-like shapes to efficiently express the stationary voiced segment, dispersion patterns of shapes such that the energy of pulse vectors output from the algebraic codebook (energy is concentrated on the positions of non-zero elements) is spread around, dispersion patterns selected from among several arbitrarily prepared dispersion pattern candidates so that a synthesized speech of high quality can be output by encoding and decoding a speech signal and repeating subjective (listening) evaluation tests of the synthesized speech or dispersion patterns created based on phonological knowledge, etc. at least one type per non-zero element (channel) in the excitation vector output from the algebraic codebook, convoluting the registered dispersion patterns and vectors generated by the algebraic codebook (made up of a small number of non-zero elements) for every channel, adding up the convolution results of respective channels and using the addition result as the stochastic excitation vector.

Moreover, especially when dispersion pattern storage section 4012 registers dispersion patterns of

a plurality of types (two or more types) per channel, methods disclosed as the methods for selecting a plurality of these dispersion patterns include: a method of actually performing encoding and decoding on all combinations of the registered dispersion patterns and "closed-loop search" a dispersion pattern corresponding to a minimum of the resulting coding distortion and a method for "open-loop search" dispersion patterns using speech-like information which is already made clear when a stochastic codebook search is performed (the speech-like information here refers to, for example, voicing strength information judged using dynamic variation information of gain codes or comparison result between gain values and a preset threshold value or voicing strength information judged using dynamic variation of linear predictive codes).

By the way, for simplicity of explanations, the following explanations will be confined to a dispersed-pulse codebook in FIG.10 characterized in that dispersion pattern storage section 4012 in the dispersed-pulse codebook in FIG.9 registers dispersion pattern of only one type per channel.

Here, the following explanation will describe stochastic codebook search processing in the case where a dispersed-pulse codebook is applied to a CELP encoder in contrast to stochastic codebook search processing in the case where an algebraic codebook is applied to a CELP encoder. First, the codebook search processing

when an algebraic codebook is used for the stochastic codebook section will be explained.

Suppose the number of non-zero elements in a vector output by the algebraic codebook is N (the number of channels of the algebraic codebook is N), a vector including only one non-zero element whose amplitude output per channel is +1 or -1 (the amplitude of elements other than non-zero elements is zero) is  $d_i$  ( $i$ : channel number:  $0 \leq i \leq N-1$ ) and the subframe length is L.

Stochastic excitation vector  $c_k$  with entry number  $k$  output by the algebraic codebook is expressed in expression 9 below:

$$C_k = \sum_{i=0}^{N-1} d_i$$

Expression 9

where:

$C_k$ : Stochastic excitation vector with entry number  $K$  according to algebraic codebook

$d_i$ : Non-zero element vector ( $d_i = \pm \delta(n - p_i)$ , where  $p_i$ : position of non-zero element)

$N$ : The number of channels of algebraic codebook (= The number of non-zero elements in stochastic excitation vector)

Then, by substituting expression 9 into expression 10, expression 11 below is obtained:

$$D_k = \frac{(v' H c_k)^2}{\|H c_k\|^2}$$

Expression 10

where:

$v^t$ : Transposition vector of  $v$  (target vector for stochastic codebook search )

$H^t$ : Transposition matrix of  $H$  (impulse response matrix of the synthesis filter)

$ck$ : Stochastic excitation vector of entry number  $k$

$$Dk = \frac{\left( v^t H \left( \sum_{i=0}^{N-1} di \right) \right)^2}{\left\| H \left( \sum_{i=0}^{N-1} di \right) \right\|^2}$$

Expression 11

10 where:

$v$ : target vector for stochastic codebook search

$H$ : Impulse response convolution matrix of the synthesis filter

$di$ : Non-zero element vector (  $di = \pm \delta(n - pi)$  , where  $pi$ : position of non-zero element)

$N$ : The number of channels of algebraic codebook (= The number of non-zero elements in stochastic excitation vector)

$$x^t = v^t H$$

20  $M = H^t H$

The processing to identify entry number  $k$  that maximizes expression 12 below obtained by arranging this expression 10 becomes stochastic codebook search processing.

$$Dk = \frac{\left( \sum_{i=0}^{N-1} x^t d_i \right)^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} d_i^t M d_j}$$

Expression 12

where,  $\mathbf{x}^t = \mathbf{v}^t \mathbf{H}$ 、 $\mathbf{M} = \mathbf{H}^t \mathbf{H}$  ( $\mathbf{v}$  is a target vector for stochastic codebook search) in expression 12. Here, when the value of expression 12 about each entry number  $k$  is calculated,  $\mathbf{x}^t = \mathbf{v}^t \mathbf{H}$  and  $\mathbf{M} = \mathbf{H}^t \mathbf{H}$  are calculated in the pre-processing stage and the calculation result is developed (stored) in memory. It is disclosed in the above papers, etc. and generally known that introducing this pre-processing makes it possible to drastically reduce the amount of computational complexity when expression 12 is calculated for every candidate entered as the stochastic excitation vector and as a result, suppress the total amount of computational complexity required for a stochastic codebook search to a small value.

Next, the stochastic codebook search processing when the dispersed-pulse codebook is used for the stochastic codebook will be explained.

Suppose the number of non-zero elements output from the algebraic codebook, which is a component of the dispersed-pulse codebook, is  $N$  ( $N$ : the number of channels of the algebraic codebook), a vector that includes only one non-zero element whose amplitude is  $+1$  or  $-1$  output for each channel (the amplitude of elements other than non-zero element is zero) is  $d_i$  ( $i$ : channel number: 0

$\leq i \leq N-1$ ), the dispersion patterns for channel number  $i$  stored in the dispersion pattern storage section is  $w_i$  and the subframe length is  $L$ . Then, stochastic excitation vector  $c_k$  of entry number  $k$  output from the dispersed-pulse codebook is given by expression 13 below:

$$C_k = \sum_{i=0}^{N-1} W_i d_i$$

Expression 13

where:

10  $C_k$ : Stochastic excitation vector of entry number  $k$  output from dispersed-pulse codebook

$W_i$ : dispersion pattern ( $w_i$ ) convolution matrix

$d_i$ : Non-zero element vector output by algebraic codebook section ( $d_i = \pm \delta(n \cdot p_i)$ , where  $p_i$ : position of non-zero element)

15  $N$ : The number of channels of algebraic codebook section

Therefore, in this case, expression 14 below is obtained by substituting expression 13 into expression 20 10.

$$D_k = \frac{\left( v^H \left( \sum_{i=0}^{N-1} W_i d_i \right) \right)^2}{\left\| H \left( \sum_{i=0}^{N-1} W_i d_i \right) \right\|^2}$$

Expression 14

where:

$v$ : target vector for stochastic codebook search

$H$ : Impulse response convolution matrix of synthesis filter

$W_i$ : Dispersion pattern ( $w_i$ ) convolution matrix

5  $d_i$ : Non-zero element vector output by typical codebook section

( $d_i = \pm \delta(n - p_i)$ , where  $p_i$ : position of non-zero element)

10  $N$ : The number of channels of algebraic codebook (= the number of non-zero elements in stochastic excitation vector)

$$H_i = HW_i$$

$$x_i^t = v^t H_i$$

$$R = H_i H_j$$

15 The processing of identifying entry number  $k$  of the stochastic excitation vector that maximizes expression 15 below obtained by arranging this expression 14 is the stochastic codebook search processing when the dispersed-pulse codebook is used.

$$20 \quad Dk = \frac{\left( \sum_{i=0}^{N-1} x_i^t d_i \right)^2}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} d_i^t R d_j}$$

Expression 15

where, in expression 15,  $x^t = v^t H_i$  (where  $H_i = HW_i$ :  $W_i$  is the dispersion pattern convolution matrix). When a value of expression 15 is calculated for each entry  
25 number  $k$ , it is possible to calculate  $H_i = HW_i$ ,  $x^t = v^t H_i$  and  $R = H_i^t H_j$  as the pre-processing and record this in memory.

calculate expression 15 for each candidate entered as a stochastic excitation vector becomes equal to the amount of computational complexity to calculate expression 12 when the algebraic codebook is used (it is obvious that expression 12 and expression 15 have the same form) and it is possible to perform a stochastic codebook search with a small amount of computational complexity even when the dispersed-pulse codebook is used.

10       The above technology shows the effects of using the dispersed-pulse codebook for the stochastic codebook section of the CELP encoder/decoder and shows that when used for the stochastic codebook section, the dispersed-pulse codebook makes it possible to perform

15       a stochastic codebook search with the same method as that when the algebraic codebook is used for the stochastic codebook section. The difference between the amount of computational complexity required for a stochastic codebook search when the algebraic codebook is used for

20       the stochastic codebook section and the amount of computational complexity required for a stochastic codebook search when the dispersed-pulse codebook is used for the stochastic codebook section corresponds to the difference between the amounts of computational

25       complexity required for the pre-processing stage of expression 12 and expression 15, that is, the difference between the amounts of computational complexity required for pre-processing ( $\mathbf{x}^t = \mathbf{v}^t \mathbf{H} \mathbf{i}$ ,  $\mathbf{M} = \mathbf{H}^t \mathbf{H}$ ) and pre-



processing ( $\mathbf{H}i = \mathbf{H}\mathbf{W}i$ ,  $\mathbf{x}^t = \mathbf{v}^t \mathbf{H}i$ ,  $\mathbf{R} = \mathbf{H}i^t \mathbf{H}j$ ).

In general, with the CELP encoder/decoder, as the bit rate decreases, the number of bits assignable to the stochastic codebook section also tends to be decreased.

5 This tendency leads to a decrease in the number of non-zero elements when a stochastic excitation vector is formed in the case where the algebraic codebook and dispersed-pulse codebook are used for the stochastic codebook section. Therefore, as the bit rate of the CELP

10 encoder/decoder decreases, the difference in the amount of computational complexity when the algebraic codebook is used and when the dispersed-pulse codebook is used decreases. However, when the bit rate is relatively high or when the amount of computational complexity needs to

15 be reduced even if the bit rate is low, the increase in the amount of computational complexity in the pre-processing stage resulting from using the dispersed-pulse codebook is not negligible.

This embodiment explains the case where in a

20 CELP-based speech encoder and speech decoder and speech encoding/decoding system using a dispersed-pulse codebook for the stochastic codebook section, the decoding side obtains synthesized speech of high quality while suppressing to a low level the increase in the

25 amount of computational complexity of the pre-processing section in the stochastic codebook search processing, which increases compared with the case where the algebraic codebook is used for the stochastic codebook

section.

More specifically, the technology according to this embodiment is intended to solve the problem above that may occur when the dispersed-pulse codebook is used  
 5 for the stochastic codebook section of the CELP encoder/decoder, and is characterized by using a dispersion pattern, which differs between the encoder and decoder. That is, this embodiment registers the above-described dispersion pattern in the dispersion  
 10 pattern storage section on the speech decoder side and generates synthesized speech of higher quality using the dispersion pattern than using the algebraic codebook.

On the other hand, the speech encoder registers a dispersion pattern, which is the simplified dispersion  
 15 pattern to be registered in the dispersion pattern storage section of the decoder (e.g., dispersion pattern selected at certain intervals or dispersion pattern truncated at a certain length) and performs a stochastic codebook search using the simplified dispersion pattern.

20 When the dispersed-pulse codebook is used for the stochastic codebook section, this allows the coding side to suppress to a small level the amount of computational complexity at the time of a stochastic codebook search in the pre-processing stage, which increases compared  
 25 to the case where the algebraic codebook is used for the stochastic codebook section and allows the decoding side to obtain a synthesized speech of high quality.

Using different dispersion patterns for the

encoder and decoder means acquiring an dispersion pattern for the encoder by modifying the prepared spreading vector (for the decoder) while reserving the characteristic.

5           Here, examples of the method for preparing a dispersion pattern for the decoder include the methods disclosed in the patent (Unexamined Japanese Patent Publication No.HEI 10-63300) applied for by the present inventor, et al., that is, a method for preparing a  
10 dispersion pattern by training of the statistic tendency of a huge number of target vectors for stochastic codebook search, a method for preparing a dispersion vector by repeating operations of encoding and decoding the actual target vector for stochastic codebook search and  
15 gradually modifying the decoded target vector in the direction in which the sum total of coding distortion generated is reduced, a method of designing based on phonological knowledge in order to achieve synthesized speech of high quality or a method of designing for the  
20 purpose of randomizing the high frequency phase component of the pulse excitation vector. All these contents are included here.

          All these dispersion patterns acquired in this way are characterized in that the amplitude of a sample close  
25 to the start sample of the dispersion pattern (forward sample) is relatively larger than the amplitude of a backward sample. Above all, the amplitude of the start sample is often the maximum of all samples in the

dispersion pattern (this is true in most cases).

The following are examples of the specific method for acquiring a dispersion pattern for the encoder by modifying the dispersion pattern for the decoder while  
5 reserving the characteristic:

1) Acquiring a dispersion pattern for the encoder by replacing the sample value of the dispersion pattern for the decoder with zero at appropriate intervals

2) Acquiring a dispersion pattern for the encoder  
10 by truncating the dispersion pattern for the decoder of a certain length at an appropriate length

3) Acquiring a dispersion pattern for the encoder by setting a threshold of amplitude beforehand and replacing a sample whose amplitude is smaller than a  
15 threshold set for the dispersion pattern for the decoder with zero

4) Acquiring a dispersion pattern for the coder by storing a sample value of the dispersion pattern for the decoder of a certain length at appropriate intervals  
20 including the start sample and replacing other sample values with zero

Here, even in the case where a few samples from the beginning of the dispersion pattern is used as in the case of the method in 1) above, for example, it is possible  
25 to acquire a new dispersion pattern for the encoder while reserving an outline (gross characteristic) of the dispersion pattern.

Furthermore, even in the case where a sample value

is replaced with zero at appropriate intervals as in the case of the method in 2) above, for example, it is possible to acquire a new dispersion pattern for the encoder while reserving an outline (gross characteristic) of the original dispersion pattern. Especially, the method in 4) above includes a restriction that the amplitude of the start sample whose amplitude is often the largest should always be saved as is, and therefore it is possible to save an outline of the original spreading vector more reliably.

Furthermore, even in the case where a sample whose amplitude is equal to or larger than a specific threshold value is saved as is and a sample whose amplitude is smaller than the specific threshold value is replaced with zero as the method in the case of 3) above, it is possible to acquire a dispersion pattern for the encoder while reserving an outline (gross characteristic) of the dispersion pattern.

The speech encoder and speech decoder according to this embodiment will be explained in detail with reference to the attached drawings below. The CELP speech encoder (FIG.11) and the CELP speech decoder (FIG.12) described in the attached drawings are characterized by using the above dispersed-pulse codebook for the stochastic codebook section of the conventional CELP speech encoder and the CELP speech decoder. Therefore, in the following explanations, it is possible to read the parts described "the stochastic

codebook", "stochastic excitation vector" and  
 "stochastic excitation vector gain" as "dispersed-pulse  
 codebook", "dispersed-pulse excitation vector" and  
 "dispersed-pulse excitation vector gain", respectively.

5 The stochastic codebook in the CELP speech encoder and  
 the CELP speech decoder has the function of storing a  
 noise codebook or fixed waveforms of a plurality of types,  
 and therefore is sometimes also called a "fixed  
 codebook".

10 In the CELP speech encoder in FIG.11, linear  
 predictive analysis section 501 performs a linear  
 predictive analysis on the input speech and calculates  
 a linear prediction coefficient first and then outputs  
 the calculated linear prediction coefficient to linear  
 15 prediction coefficient encoding section 502. Then,  
 linear prediction coefficient encoding section 502  
 performs encoding (vector quantization) on the linear  
 prediction coefficient and outputs the quantization  
 index (hereinafter referred to as "linear predictive  
 20 code") obtained by vector quantization to code output  
 section 513 and linear predictive code decoding section  
 503.

Then, linear predictive code decoding section 503  
 performs decoding (inverse-quantization) on the linear  
 25 predictive code obtained by linear prediction  
 coefficient encoding section 502 and outputs to  
 synthesis filter 504. Synthesis filter 504 constitutes  
 a synthesis filter having the all-pole model structure

based on the decoding linear predictive code obtained from linear predictive code decoding section 503.

Then, vector adder 511 adds up a vector obtained by multiplying the adaptive excitation vector selected from adaptive codebook 506 by adaptive excitation vector gain 509 and a vector obtained by multiplying the stochastic excitation vector selected from dispersed-pulse codebook 507 by stochastic excitation vector gain 510 to generate an excitation vector. Then, distortion calculation section 505 calculates distortion between the output vector when synthesis filter 504 is excited by the excitation vector and the input speech according to expression 16 below and outputs distortion ER to code identification section 512.

$$ER = \|u - (g_a Hp + g_c Hc)\|^2$$

Expression 16

where:

u: Input speech (vector)  
 H: Impulse response matrix of synthesis filter  
 p: Adaptive excitation vector  
 c: Stochastic excitation vector  
 $g_a$ : Adaptive excitation vector gain  
 $g_c$ : Stochastic excitation vector gain

In expression 16, u denotes an input speech vector inside the frame being processed, H denotes an impulse response matrix of synthesis filter,  $g_a$  denotes an adaptive excitation vector gain,  $g_c$  denotes a stochastic excitation vector gain, p denotes an adaptive excitation

vector and  $c$  denotes a stochastic excitation vector.

Here, adaptive codebook 506 is a buffer (dynamic memory) that stores excitation vectors corresponding a several number of past frames and the adaptive excitation  
 5 vector selected from adaptive codebook 506 above is used to express the periodic component in the linear predictive residual vector obtained by passing the input speech through the inverse-filter of the synthesis filter.

10 On the other hand, the excitation vector selected from dispersed-pulse codebook 507 is used to express the non-periodic (the component obtained by removing periodic component (adaptive excitation vector component) from the linear predictive residual vector)  
 15 newly added to the linear predictive residual vector in the frame actually being processed.

Adaptive excitation vector gain multiplication section 509 and stochastic excitation vector gain multiplication section 510 have the function of  
 20 multiplying the adaptive excitation vector selected from adaptive codebook 506 and stochastic excitation vector selected from dispersed-pulse codebook 507 by the adaptive excitation vector gain and stochastic excitation vector gain read from gain codebook 508. Gain  
 25 codebook 508 is a static memory that stores a plurality of types of sets of an adaptive excitation vector gain to be multiplied on the adaptive excitation vector and stochastic excitation vector gain to be multiplied on



the stochastic excitation vector.

Code identification section 512 selects an optimal combination of indices of the three codebooks above (adaptive codebook, dispersed-pulse codebook, gain codebook) that minimizes distortion ER of expression 16 calculated by distortion calculation section 505. Then, distortion identification section 512 outputs the indices of their respective codebooks selected when the above distortion reaches a minimum to code output section 513 as adaptive excitation vector code, stochastic excitation vector code and gain code, respectively.

Finally, code output section 513 compiles the linear predictive code obtained from linear prediction coefficient encoding section 502 and the adaptive excitation vector code, stochastic excitation vector code and gain code identified by code identification section 512 into a code (bit information) that expresses the input speech inside the frame actually being processed and outputs this code to the decoder side.

By the way, code identification section 512 sometimes identifies an adaptive excitation vector code, stochastic excitation vector code and gain code on a "subframe" basis, where "subframe" is a subdivision of the processing frame. However, no distinction will be made between a frame and a subframe (will be commonly referred to as "frame") in the following explanations of the present description.

Then, an outline of the CELP speech decoder will

be explained using FIG.12.

In the CELP decoder in FIG.12, code input section 601 receives a code (bit information to reconstruct a speech signal on a (sub) frame basis) identified and transmitted from the CELP speech encoder (FIG.11) and de-multiplexes the received code into 4 types of code: a linear predictive code, adaptive excitation vector code, stochastic excitation vector code and gain code. Then, code input section 601 outputs the linear predictive code to linear prediction coefficient decoding section 602, the adaptive excitation vector code to adaptive codebook 603, the stochastic excitation vector code to dispersed-pulse codebook 604 and the gain code to gain codebook 605.

Then, linear prediction coefficient decoding section 602 decodes the linear predictive code input from code input section 601, obtains a decoded linear predictive coefficients and outputs this decoded linear predictive coefficients to synthesis filter 609.

Synthesis filter 609 constructs a synthesis filter having the all-pole model structure based on the decoding linear predictive code obtained from linear predictive code decoding section 602. On the other hand, adaptive codebook 603 outputs an adaptive excitation vector corresponding to the adaptive excitation vector code input from code input section 601. Dispersed-pulse codebook 604 outputs a stochastic excitation vector corresponding to the stochastic excitation vector code

input from code input section 601. Gain codebook 605 reads an adaptive excitation gain and stochastic excitation gain corresponding to the gain code input from code input section 601 and outputs these gains to

5 adaptive excitation vector gain multiplication section 606 and stochastic excitation vector gain multiplication section 607, respectively.

Then, adaptive excitation vector gain multiplication section 606 multiplies the adaptive excitation vector output from adaptive codebook 603 by

10 the adaptive excitation vector gain output from gain codebook 605 and stochastic excitation vector gain multiplication section 607 multiplies the stochastic excitation vector output from dispersed-pulse codebook 604 by the stochastic excitation vector gain output from gain codebook 605. Then, vector addition section 608 adds up the respective output vectors of adaptive excitation vector gain multiplication section 606 and stochastic excitation vector gain multiplication

15 section 607 to generate an excitation vector. Then, synthesis filter 609 is excited by this excitation vector and a synthesized speech of the received frame section is output.

20

It is important to suppress distortion ER of

25 expression 16 to a small value in order to obtain a synthesized speech of high quality in such a CELP-based speech encoder/speech decoder. To do this, it is desirable to identify the best combination of an adaptive

excitation vector code, stochastic excitation vector code and gain code in closed-loop fashion so that ER of expression 16 is minimized. However, since attempting to identify distortion ER of expression 16 in the

5 closed-loop fashion leads to an excessively large amount of computational complexity, it is a general practice to identify the above 3 types of code in the open-loop fashion.

More specifically, an adaptive codebook search is

10 performed first. Here, the adaptive codebook search processing refers to processing of vector quantization of the periodic component in a predictive residual vector obtained by passing the input speech through the

inverse-filter by the adaptive excitation vector output

15 from the adaptive codebook that stores excitation vectors of the past several frames. Then, the adaptive codebook search processing identifies the entry number of the adaptive excitation vector having a periodic component close to the periodic component within the

20 linear predictive residual vector as the adaptive excitation vector code. At the same time, the adaptive codebook search temporarily ascertains an ideal adaptive excitation vector gain.

Then, a stochastic codebook search (corresponding

25 to dispersed-pulse codebook search in this embodiment) is performed. The dispersed-pulse codebook search refers to processing of vector quantization of the linear predictive residual vector of the frame being processed

with the periodic component removed, that is, the component obtained by subtracting the adaptive excitation vector component from the linear predictive residual vector (hereinafter also referred to as "target vector for stochastic codebook search") using a plurality of stochastic excitation vector candidates generated from the dispersed-pulse codebook. Then, this dispersed-pulse codebook search processing identifies the entry number of the stochastic excitation vector that performs encoding of the target vector for stochastic codebook search with least distortion as the stochastic excitation vector code. At the same time, the dispersed-pulse codebook search temporarily ascertains an ideal stochastic excitation vector gain.

Finally, a gain codebook search is performed. The gain codebook search is processing of encoding (vector quantization) on a vector made up of 2 elements of the ideal adaptive gain temporarily obtained during the adaptive codebook search and the ideal stochastic gain temporarily obtained during the dispersed-pulse codebook search so that distortion with respect to a gain candidate vector (vector candidate made up of 2 elements of the adaptive excitation vector gain candidate and stochastic excitation vector gain candidate) stored in the gain codebook reaches a minimum. Then, the entry number of the gain candidate vector selected here is output to the code output section as the gain code.

Here, of the general code search processing above

in the CELP speech encoder, the dispersed-pulse codebook search processing (processing of identifying a stochastic excitation vector code after identifying an adaptive excitation vector code) will be explained in further detail below.

As explained above, a linear predictive code and adaptive excitation vector code are already identified when a dispersed-pulse codebook search is performed in a general CELP encoder. Here, suppose an impulse response matrix of a synthesis filter made up of an already identified linear predictive code is  $H$ , an adaptive excitation vector corresponding to an adaptive excitation vector code is  $p$  and an ideal adaptive excitation vector gain (provisional value) determined simultaneously with the identification of the adaptive excitation vector code is  $g_a$ . Then, distortion  $ER$  of expression 16 is modified into expression 17 below.

$$ER_k = \|v - g_c H c_k\|^2$$

Expression 17

where:

$v$ : Target vector for stochastic codebook search  
(where,  $v = u - g_a H p$ )

$g_c$ : Stochastic excitation vector gain

$H$ : Impulse response matrix of a synthesis filter

$c_k$ : Stochastic excitation vector ( $k$ : entry number)

Here, vector  $v$  in expression 17 is the target vector for stochastic codebook search of expression 18 below using input speech signal  $u$  in the processing frame,

impulse response matrix  $H$  (determined) of the synthesis filter, adaptive excitation vector  $p$  (determined) and ideal adaptive excitation vector gain  $g_a$  (provisional value).

$$5 \quad v = u - g_a H p$$

Expression 18

where:

$u$ : Input speech (vector)

$g_a$ : Adaptive excitation vector gain (provisional  
10 value)

$H$ : Impulse response matrix of a synthesis filter

$p$ : Stochastic excitation vector

By the way, the stochastic excitation vector is  
15 expressed as " $c$ " in expression 16, while the stochastic  
excitation vector is expressed as " $ck$ " in expression 17.  
This is because expression 16 does not explicitly  
indicate the difference of the entry number ( $k$ ) of the  
stochastic excitation vector, whereas expression 17  
20 explicitly indicates the entry number. Despite the  
difference in expression, both are the same in meaning.

Therefore, the dispersed-pulse codebook search  
means the processing of determining entry number  $k$  of  
stochastic excitation vector  $ck$  that minimizes  
25 distortion  $ER_k$  of expression 17. Moreover, when entry  
number  $k$  of stochastic excitation vector  $ck$  that  
minimizes distortion  $ER_k$  of expression 17 is identified,  
stochastic excitation gain  $g_c$  is assumed to be able to

take an arbitrary value. Therefore, the processing of determining the entry number that minimizes distortion of expression 17 can be replaced with the processing of identifying entry number  $k$  of stochastic excitation vector  $c_k$  that maximizes  $D_k$  of expression 10 above.

Then, the dispersed-pulse codebook search is carried out in 2 stages: distortion calculation section 505 calculates  $D_k$  of expression 10 for every entry number  $k$  of stochastic excitation vector  $c_k$ , outputs the value to code identification section 512 and code identification section 512 compares the values, large and small, in expression 10 for every entry number  $k$ , determines entry number  $k$  when the value reaches a maximum as the stochastic excitation vector code and outputs to code output section 513.

The operations of the speech encoder and speech decoder according to this embodiment will be explained below.

FIG.13A shows a configuration of dispersed-pulse codebook 507 in the speech encoder shown in FIG.11 and FIG.13B shows a configuration of dispersed-pulse codebook 604 in the speech decoder shown in FIG.12. The difference in configuration between dispersed-pulse codebook 507 shown in FIG.13A and dispersed-pulse codebook 604 shown in FIG.13B is the difference in the shape of dispersion patterns registered in the dispersion pattern storage section.

In the case of the speech decoder in FIG.13B,



dispersion pattern storage section 4012 registers one type per channel of any one of (1) dispersion pattern of a shape resulting from statistical training of shapes of a huge number of target vectors for stochastic codebook search, contained in a target vector for stochastic codebook search, (2) dispersion pattern of a random-like shape to efficiently express unvoiced consonant segments and noise-like segments, (3) dispersion pattern of a pulse-like shape to efficiently express stationary voiced segments, (4) dispersion pattern of a shape that gives an effect of spreading around the energy (the energy is concentrated on the positions of non-zero elements) of an excitation vector output from the algebraic codebook, (5) dispersion pattern selected from among several arbitrarily prepared dispersion pattern candidates by repeating encoding and decoding of the speech signal and an subjective (listening) evaluation of the synthesized speech so that synthesized speech of high quality can be output and (6) dispersion pattern created based on phonological knowledge.

On the other hand, dispersion pattern storage section 4012 in the speech encoder in FIG.13A registers dispersion patterns obtained by replacing dispersion patterns registered in dispersion pattern storage section 4012 in the speech decoder in FIG.13B with zero for every other sample.

Then, the CELP speech encoder/speech decoder in the

above configuration encodes/decodes the speech signal using the same method as described above without being aware that different dispersion patterns are registered in the encoder and decoder.

5           The encoder can reduce the amount of computational complexity of pre-processing during a stochastic codebook search when the dispersed-pulse codebook is used for the stochastic codebook section (can reduce by half the amount of computational complexity of  $H_i = H_t W_i$  and  $X_{it} = v_t H_i$ ), while the decoder can spread around the energy concentrated on the positions of non-zero elements by convoluted conventional dispersion patterns on pulse vectors, making it possible to improve the quality of a synthesized speech.

10           As shown in FIG.13A and FIG.13B, this embodiment describes the case where the speech encoder uses dispersion patterns obtained by replacing dispersion patterns used by the speech decoder with zero every other sample. However, this embodiment is also directly applicable to a case where the speech encoder uses dispersion patterns obtained by replacing dispersion pattern elements used by the speech decoder with zero every  $N$  ( $N \geq 1$ ) samples, and it is possible to attain similar action in that case, too.

25           Furthermore, this embodiment describes the case where the dispersion pattern storage section registers dispersion patterns of one type per channel, but the present invention is also applicable to a CELP speech

encoder/decoder that uses the dispersed-pulse codebook characterized by registering dispersion patterns of 2 or more types per channel and selecting and using a dispersion pattern for the stochastic codebook section,  
 5 and it is possible to attain similar actions and effects in that case, too.

Furthermore, this embodiment describes the case where the dispersed-pulse codebook use an algebraic codebook that outputs a vector including 3 non-zero  
 10 elements, but this embodiment is also applicable to a case where the vector output by the algebraic codebook section includes  $M$  ( $M \geq 1$ ) non-zero elements, and it is possible to attain similar actions and effects in that case, too.

15 Furthermore, this embodiment describes the case where an algebraic codebook is used as the codebook for generating a pulse vector made up of a small number of non-zero elements, but this embodiment is also applicable to a case where other codebooks such as  
 20 multi-pulse codebook or regular pulse codebook are used as the codebooks for generating the relevant pulse vector, and it is possible to attain similar actions and effects in that case, too.

Then, FIG.14A shows a configuration of the  
 25 dispersed-pulse codebook in the speech encoder in FIG.11 and FIG.14B shows a configuration of the dispersed-pulse codebook in the speech decoder in FIG.12.

The difference in configuration between the

dispersed-pulse codebook shown in FIG.14A and the dispersed-pulse codebook shown in FIG.14B is the difference in the length of dispersion patterns registered in the dispersion pattern storage section.

- 5 In the case of the speech decoder in FIG.14B, dispersion pattern storage section 4012 registers one type per channel of any one of (1) dispersion pattern of a shape resulting from statistical training of shapes based on a huge number of target vectors for stochastic codebook
- 10 search, (2) dispersion pattern of a random-like shape to efficiently express unvoiced consonant segments and noise-like segments, (3) dispersion pattern of a pulse-like shape to efficiently express stationary voiced segments, (4) dispersion pattern of a shape that
- 15 gives an effect of spreading around the energy (the energy is concentrated on the positions of non-zero elements) of an excitation vector output from the algebraic codebook, (5) dispersion pattern selected from among several arbitrarily prepared dispersion pattern
- 20 candidates by repeating encoding and decoding of the speech signal and subjective(listening) evaluation of the synthesized speech so that synthesized speech of high quality can be output and (6) dispersion pattern created based on phonological knowledge.

- 25 On the other hand, dispersion pattern storage section 4012 in the speech encoder in FIG.14A registers dispersion patterns obtained by truncating dispersion patterns registered in the dispersion pattern storage

section in the speech decoder in FIG.14B at a half length.

Then, the CELP speech encoder/speech decoder in the above configurations encodes/decodes the speech signal using the same method as described above without being  
 5 aware that different dispersion patterns are registered in the encoder and decoder.

The coder can reduce the amount of computational complexity of pre-processing during a stochastic codebook search when the dispersed-pulse codebook is  
 10 used for the stochastic codebook section (can reduce by half the amount of computational complexities of  $H_i = H_t W_i$  and  $X_{it} = v_t H_i$ ), while the decoder uses the same conventional dispersion patterns, making it possible to improve the quality of a synthesized speech.

15 As shown in FIG.14A and FIG.14B, this embodiment describes the case where the speech encoder uses dispersion patterns obtained by truncating dispersion patterns used by the speech decoder at a half length. However, when dispersion patterns used by the speech  
 20 decoder are truncated at a shorter length  $N$  ( $N \geq 1$ ), this embodiment provides an effect that it is possible to further reduce the amount of computational complexity of pre-processing during a stochastic codebook search. However, the case where dispersion patterns used by the  
 25 speech encoder are truncated at a length of 1 corresponds to the speech encoder that uses no dispersion pattern (dispersion patterns are applied to the speech decoder).

Furthermore, this embodiment describes the case

where the dispersion pattern storage section registers dispersion patterns of one type per channel, but the present invention is also applicable to a speech encoder/decoder that uses the dispersed-pulse codebook characterized by registering dispersion patterns of 2 or more types per channel and selecting and using a dispersion pattern for the stochastic codebook section, and it is possible to attain similar actions and effects in that case, too.

Furthermore, this embodiment describes the case where the dispersed-pulse codebook uses an algebraic codebook that outputs a vector including 3 non-zero elements, but this embodiment is also applicable to a case where the vector output by the algebraic codebook section includes  $M$  ( $M \geq 1$ ) non-zero elements, and it is possible to attain similar actions and effects in that case, too.

Furthermore, this embodiment describes the case where the speech encoder uses dispersion patterns obtained by truncating the dispersion patterns used by the speech decoder at a half length, but it is also possible for the speech encoder to truncate the dispersion patterns used by the speech decoder at a length of  $N$  ( $N \geq 1$ ) and further replace the truncated dispersion patterns with zero every  $M$  ( $M \geq 1$ ) samples, and it is possible to further reduce the amount of computational complexity for the stochastic codebook search.

Thus, according to this embodiment, the CELP-based speech encoder, decoder or speech encoding/decoding system using the dispersed-pulse codebook for the stochastic codebook section registers fixed waveforms frequently included in target vectors for stochastic codebook search acquired by statistical training as dispersion vectors, convolutes (reflects) these dispersion patterns on pulse vectors, and can thereby use stochastic excitation vectors, which is closer to the actual target vectors for stochastic codebook search, providing advantageous effects such as allowing the decoding side to improve the quality of synthesized speech while allowing the encoding side to suppress the amount of computational complexity for the stochastic codebook search, which is sometimes problematic when the dispersed-pulse codebook is used for the stochastic codebook section, to a lower level than conventional arts.

This embodiment can also attain similar actions and effects in the case where other codebooks such as multi-pulse codebook or regular pulse codebook, etc. are used as the codebooks for generating pulse vectors made up of a small number of non-zero elements.

The speech encoding/decoding according to Embodiments 1 to 3 above are described as the speech encoder/speech decoder, but this speech encoding/decoding can also be implemented by software. For example, it is also possible to store a program of

speech encoding/decoding described above in ROM and implement encoding/decoding under the instructions from a CPU according to the program. It is further possible to store the program, adaptive codebook and stochastic codebook (dispersed-pulse codebook) in a computer-  
5 readable recording medium, record the program, adaptive codebook and stochastic codebook (dispersed-pulse codebook) of this recording medium in RAM of the computer and implement encoding/decoding according to the program.  
10 In this case, it is also possible to attain similar actions and effects to those in Embodiments 1 to 3 above. Moreover, it is also possible to download the program in Embodiments 1 to 3 above through a communication terminal and allow this communication terminal to run  
15 the program.

Embodiments 1 to 3 can be implemented individually or combined with one another.

This application is based on the Japanese Patent Application No.HEI 11-235050 filed on August 23, 1999,  
20 the Japanese Patent Application No.HEI 11-236728 filed on August 24, 1999 and the Japanese Patent Application No.HEI 11-248363 filed on September 2, 1999, entire content of which is expressly incorporated by reference herein.

25

#### Industrial Applicability

The present invention is applicable to a base station apparatus or communication terminal apparatus



in a digital communication system.

Office  
19

09/807427

20 APR 2001

Amended Claims

1. A speech encoder comprising:

LPC synthesizing means for obtaining a synthesized  
speech by filtering adaptive excitation vector and  
5 stochastic excitation vector stored in an adaptive  
codebook and stochastic codebook using an LPC  
coefficients obtained from an input speech;

gain calculating means for calculating gains of  
said adaptive excitation vector and said stochastic  
10 excitation vector and searching code of the adaptive  
excitation vector and searching code of the stochastic  
excitation vector using coding distortion between said  
input speech and said synthesized speech obtained using  
said gains; and

15 parameter coding means for performing predictive  
coding of gains using the adaptive excitation vector and  
stochastic excitation vector corresponding to the codes  
obtained, wherein said parameter coding means comprises  
prediction coefficient adjusting means for adjusting  
20 one or more prediction coefficients used for said  
predictive coding according to one or more states of a  
previous subframe.

2. The speech encoder according to claim 1, wherein  
when one ore more states of a previous subframes are an  
25 extremely large value or an extremely small value, said  
prediction coefficient adjusting means adjusts said  
prediction coefficients so as to reduce the influence  
thereof.

3. The speech encoder according to claim 1, wherein said parameter coding means has a codebook including gain vectors of the adaptive excitation vectors, logarithmic gain vectors of the stochastic excitation vectors and  
5 coefficients for adjusting the prediction coefficient.

4. The speech encoder according to claim 3, wherein in predicting coding when a product sum between states and prediction coefficients are calculated, prediction coefficient adjustment coefficients corresponding to  
10 the states are multiplied.

5. The speech encoder according to claim 1, further comprising storing means for storing said adaptive excitation vector, said stochastic excitation vector and prediction coefficient adjustment coefficient in  
15 accordance with each state.

6. (Amended) The speech encoder according to claim 5, wherein when said adaptive excitation vector and said stochastic excitation vector stored in said storing means are updated, said prediction coefficient  
20 adjustment coefficients are also updated.

7. A CELP-based speech encoder that performs encoding by decomposing one frame into a plurality of subframes, comprising:

LPC synthesizing means for obtaining a synthesized  
25 speech by filtering adaptive excitation vector and stochastic excitation vector stored in an adaptive codebook and stochastic codebook using LPC coefficients obtained from an input speech;

gain calculating means for calculating gains of said adaptive excitation vector and said stochastic excitation vector; and

parameter coding means for performing vector  
 5 quantization of the adaptive excitation vector and stochastic excitation vector obtained using coding distortion between said input speech and said synthesized speech and said gains, and further comprising:

10 pitch analyzing means for performing pitch analyses of a plurality of subframes in the frame respectively, before performing an adaptive codebook search for the first subframe, finding a correlation value and calculating a value most approximate to the  
 15 pitch period using said correlation values.

8. The speech encoder according to claim 7, further comprising search range setting means for determining a lag search range of a plurality of subframes based on the correlation value and the value most approximate to  
 20 the pitch period obtained by said pitch analyzing means.

9. The speech encoder according to claim 8, wherein said search range setting means determines a provisional pitch that becomes the center of the search range using the correlation values and the values most approximate  
 25 to the pitch period obtained by said pitch analyzing means.

10. The speech encoder according to claim 9, wherein the search range setting means sets a lag search section

in a specified range around the provisional pitch.

11. The speech encoder according to claim 8, wherein the search range setting means sets a lag search section by reducing the number of candidates for shortpitch  
5 periods.

12. The speech encoder according to claim 8, wherein the search range setting means performs a lag search within a set range during an adaptive codebook search.

13. A computer-readable recording media storing a  
10 speech encoding program, an adaptive codebook storing past used for synthesising excitation vector signals and a stochastic codebook storing a plurality of stochastic excitation vectors, said speech encoding program comprising the steps of:

15 obtaining a synthesized speech by filtering adaptive excitation vector and by filtering stochastic excitation vector stored in said adaptive codebook and in said stochastic codebook using decoded LPC coefficients obtained from an input speech;

20 calculating gains of said adaptive excitation vector and said stochastic excitation vector;

performing vector quantization on the adaptive excitation vector and stochastic excitation vector determined using coding distortion between said input  
25 speech and said synthesized speech, and said gains, wherein said vector quantization step further comprising the steps of:

determining a quantization target vector based on

coding distortion between a plurality of quantization target vector and prediction coefficient used for predictive coding; and

adjusting said prediction coefficients according to one or more states of on or more previous subframe.

14. A computer-readable recording medium storing a speech encoding program, an adaptive codebook storing past used for synthesising excitation vector signal and a stochastic codebook storing a plurality of stochastic excitation vector, said speech coding program comprising the steps of:

obtaining a synthesized speech by filtering adaptive excitation vector and stochastic excitation vector stored in said adaptive codebook and said stochastic codebook using decoded LPC coefficients obtained from an input speech;

calculating gains of said adaptive excitation vector and said stochastic excitation vector;

performing vector quantization on the adaptive excitation vector and stochastic excitation vector determined using coding distortion between said input speech and said synthesized speech; and

calculating correlation values by performing pitch analyses of a plurality of subframes in the processing frame before performing an adaptive codebook search of the first subframe and calculating a value most approximate to the pitch period using said correlation values.

15. (Added) A speech encoder comprising a dispersed-pulse codebook that generates a vector by convoluting a vector containing one or more one non-zero elements (elements other than non-zero elements have values of zero) and a fixed waveform called a "dispersion pattern", wherein said dispersed-pulse codebook has a configuration different from that of the dispersed-pulse codebook on the speech decoder side.

16. (Added) The speech encoder according to claim 15, wherein the dispersion pattern storage section, which is a component of the dispersed-pulse codebook, stores dispersion patterns different from those stored in the dispersion pattern storage section on the speech decoder side.

17. (Added) The speech encoder according to claim 16, wherein the dispersion pattern storage section stores dispersion patterns obtained by simplifying and selecting dispersion patterns stored in the dispersion pattern storage section on the speech decoder side.

18. (Added) The speech encoder according to claim 16, wherein the dispersion pattern storage section stores dispersion patterns obtained by replacing components of dispersion patterns stored in the dispersion pattern storage section on the speech decoder side with zero at certain intervals.

19. (Added) The speech encoder according to claim 16, wherein the dispersion pattern storage section stores dispersion patterns obtained by replacing

components of dispersion patterns stored in the dispersion pattern storage section on the speech decoder side with zero for every N samples (N: natural number).

20. (Added) The speech encoder according to claim 5 19, wherein the dispersion pattern storage section stores dispersion patterns obtained by replacing components of dispersion patterns stored in the dispersion pattern storage section on the speech decoder side with zero for every 1 sample.

10 21. (Added) The speech encoder according to claim 16, wherein the dispersion pattern storage section stores dispersion patterns obtained by truncating components of dispersion patterns stored in the dispersion pattern storage section on the speech decoder side at an appropriate length.

15 22. (Added) The speech encoder according to claim 16, wherein the dispersion pattern storage section stores dispersion patterns obtained by truncating components of dispersion patterns stored in the dispersion pattern storage section on the speech decoder side at a length of N samples (N: natural number).

20 23. (Added) The speech encoder according to claim 16, wherein the dispersion pattern storage section stores dispersion patterns obtained by truncating components of dispersion patterns stored in the dispersion pattern storage section on the speech decoder side at a half length.

24. (Added) A speech decoder that decodes a speech



signal having a speech code generated by the speech encoder according to claim 15.

25. (Added) A signal processing processor containing a software program that implements the speech  
5 encoder according to claim 15.

26. (Added) A signal processing processor containing a software program that implements the speech decoder according to claim 24.

27. (Added) A speech encoding/decoding system  
10 comprising a speech encoder and/or a speech decoder each having a dispersed-pulse codebook in configuration different from each other.

28. (Added) The speech encoding/decoding system according to claim 27, wherein the difference in the  
15 configuration of the dispersed-pulse codebook between the speech encoder and the speech decoder lies in the shape of dispersion patterns stored in the respective dispersed-pulse codebooks.

29. (Added) The speech encoding/decoding system  
20 according to claim 28, wherein the shapes of dispersion patterns of the speech encoder are obtained by simplifying the shape of dispersion patterns of the speech decoder.

30. (Added) The speech encoding/decoding system  
25 according to claim 27, wherein the shapes of dispersion patterns of the speech encoder are obtained by replacing components of the dispersion patterns of the speech decoder with zero at appropriate intervals.

31. (Added) The speech encoding/decoding system according to claim 27, wherein the shapes of dispersion patterns of the speech encoder are obtained by replacing components of the dispersion patterns of the speech  
5 decoder with zero every N samples (N: natural number).

32. (Added) The speech encoding/decoding system according to claim 31, wherein the shapes of dispersion patterns of the speech encoder are obtained by replacing components of the dispersion patterns of the speech  
10 decoder with zero every 1 sample.

33. (Added) The speech encoding/decoding system according to claim 27, wherein the shapes of dispersion patterns of the speech encoder are obtained by truncating components of the dispersion patterns of the speech  
15 decoder at an appropriate length.

34. (Added) The speech encoding/decoding system according to claim 27, wherein the shapes of dispersion patterns of the speech encoder are obtained by truncating components of the dispersion patterns of the speech  
20 decoder at a length of N samples (N: natural number).

35. (Added) The speech encoding/decoding system according to claim 27, wherein the shapes of dispersion patterns of the speech encoder are obtained by truncating components of the dispersion patterns of the speech  
25 decoder at a half length.

36. (Added) A communication base station equipped with the signal processing processor according to claim 25.

37. (Added) A communication terminal equipped with the signal processing processor according to claim 25.

38. (Added) A radio communication system that connects the communication base station according to  
5 claim 36 with the communication terminal according to claim 37 via a radio network.

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2210  
2211  
2212  
2213  
2214  
2215  
2216  
2217  
2218  
2219  
2220  
2221  
2222  
2223  
2

## ABSTRACT

A vector codebook 1094 storing a plurality of typical samples of quantization target vectors is created. Each vector consists of three elements, which are values  
 5 corresponding to logarithmic values of an AC gain and SC gain and an adjustment coefficient of SC prediction coefficient. Prediction coefficient storage section 1095 stores coefficients to perform predictive coding. These coefficients are MA prediction coefficients and  
 10 a number of coefficients corresponding to the degree of prediction, of two types, AC and SC, are stored.

Parameter calculation section 1091 calculates parameters necessary for distance calculations from the input perceptual weighted input speech, perceptual  
 15 weighted LPC synthesis of adaptive code vector, perceptual weighted LPC synthesis of stochastic code vector, further decoded vectors (AC, SC, adjustment coefficient) stored in decoded vector storage section 1096 and prediction coefficients (AC, SC) stored in  
 20 prediction coefficient storage section 1095.

1 / 1 4

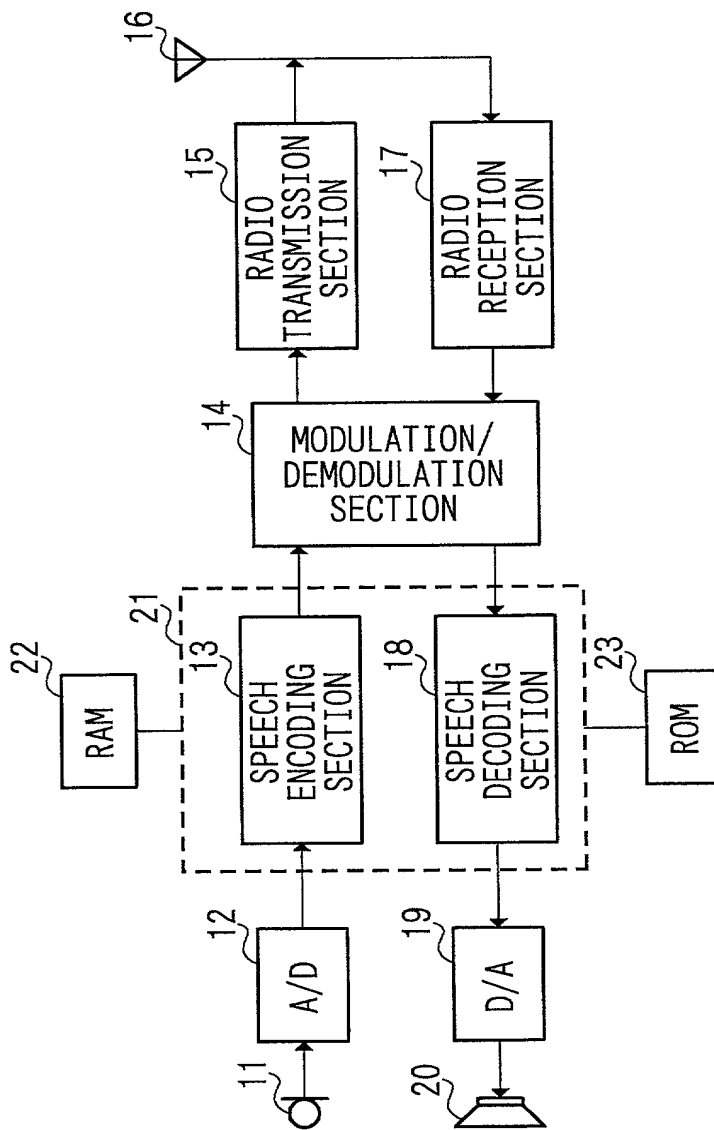


FIG. 1

2 / 1 4

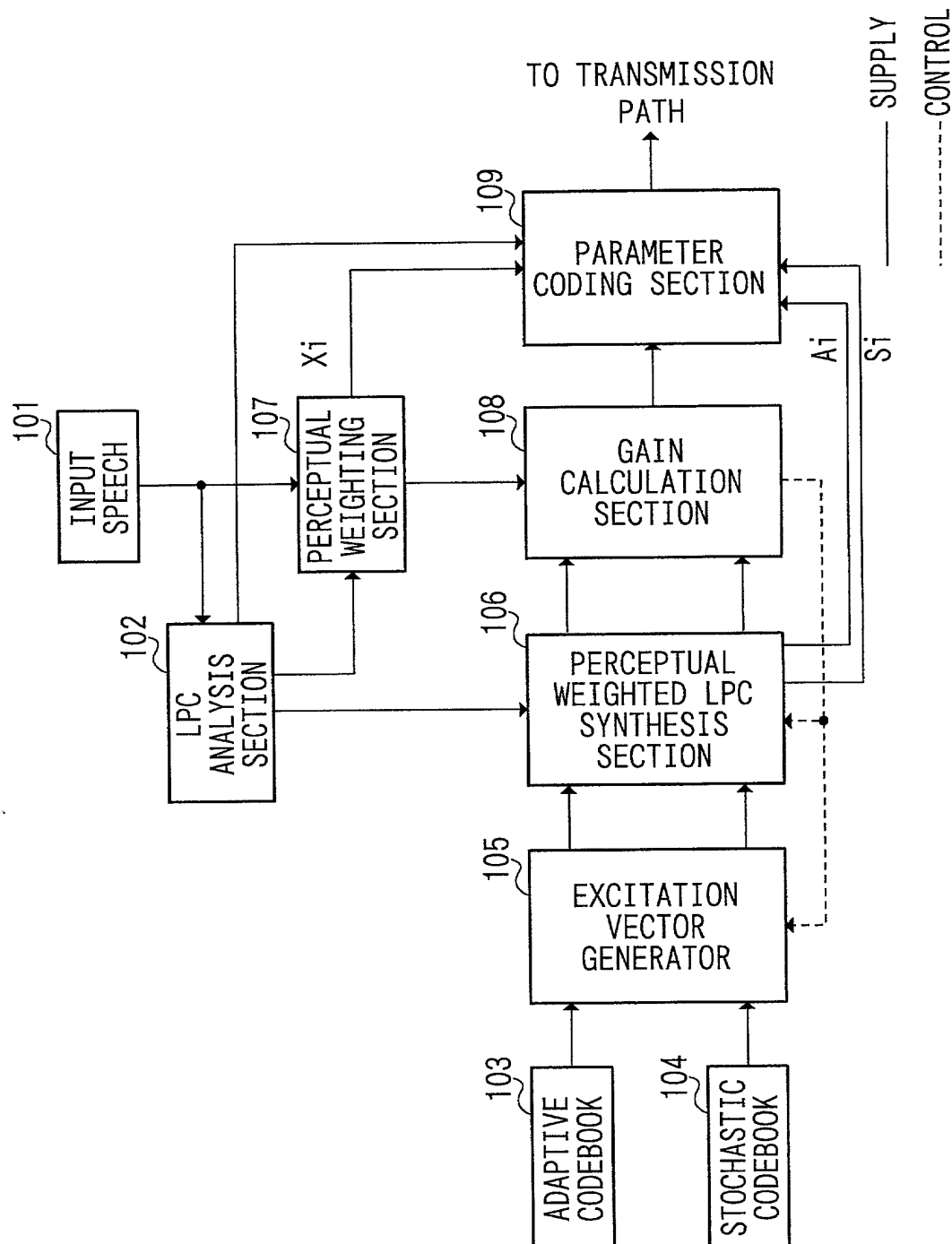


FIG. 2

3 / 1 4

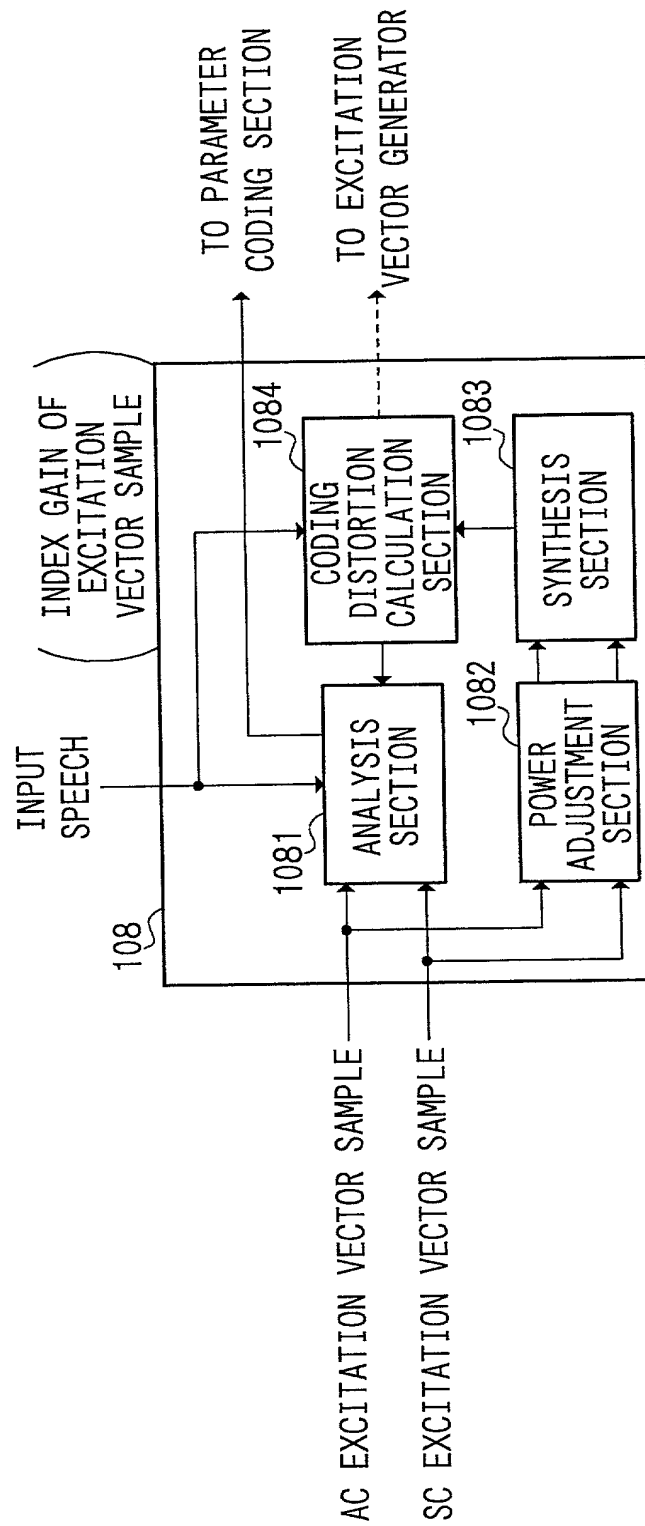


FIG. 3

4 / 1 4

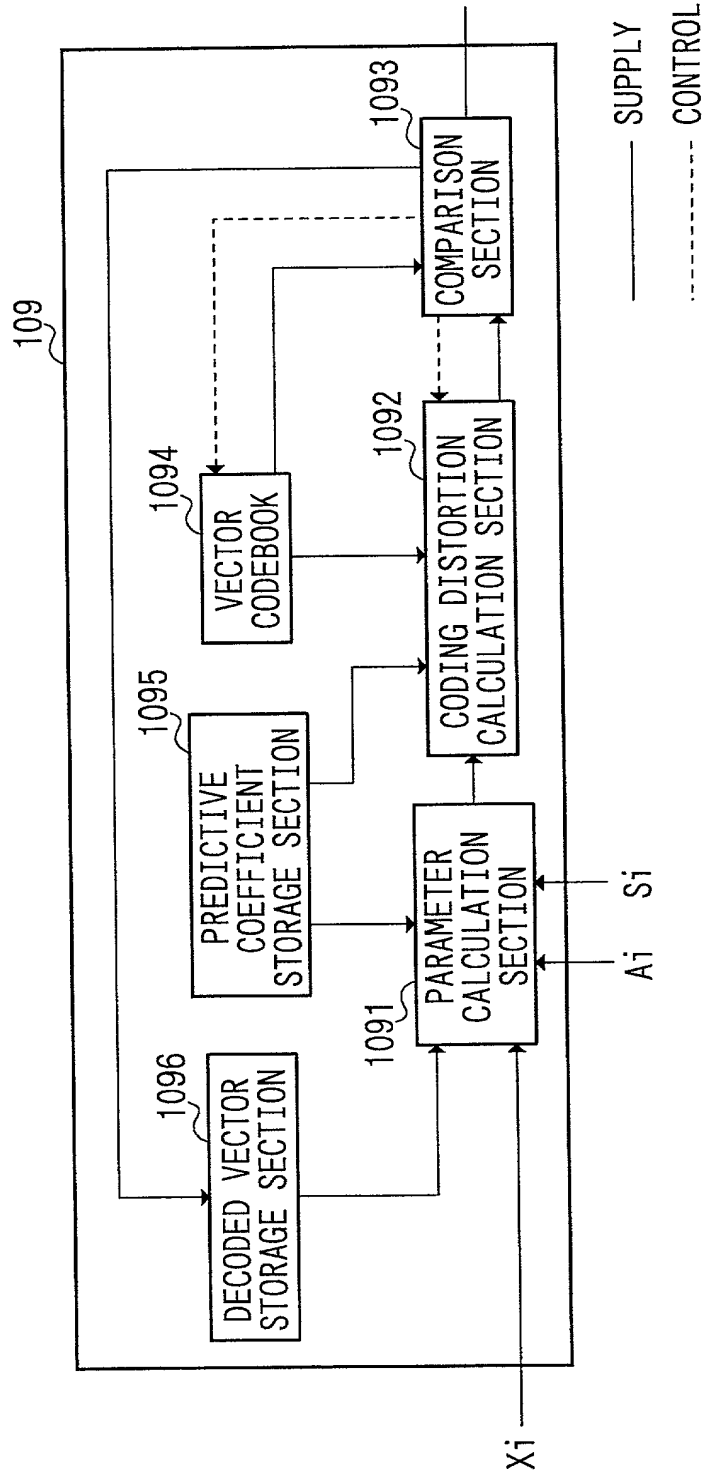


FIG. 4



5 / 1 4

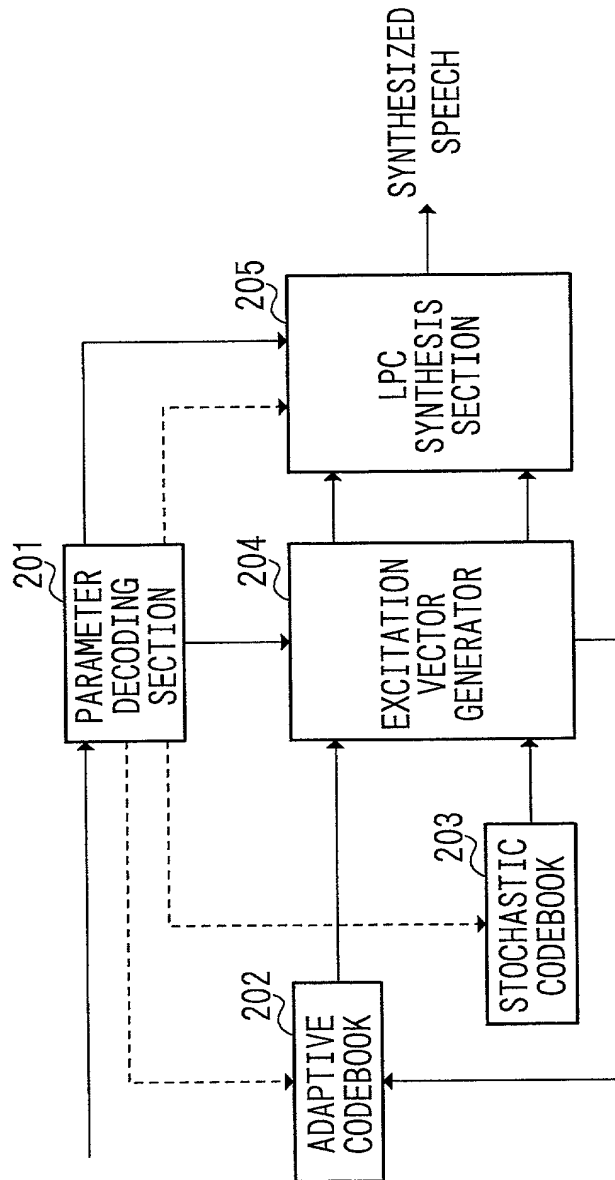


FIG. 5

6 / 1 4

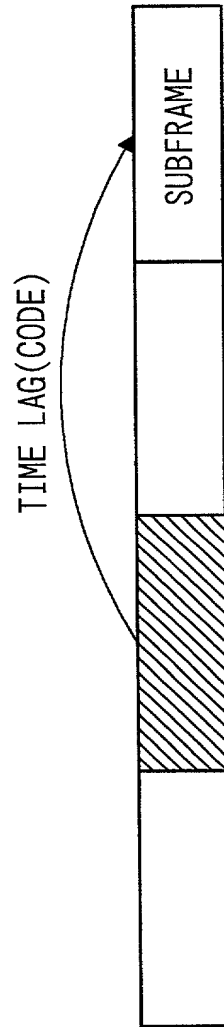


FIG. 6

7 / 1 4

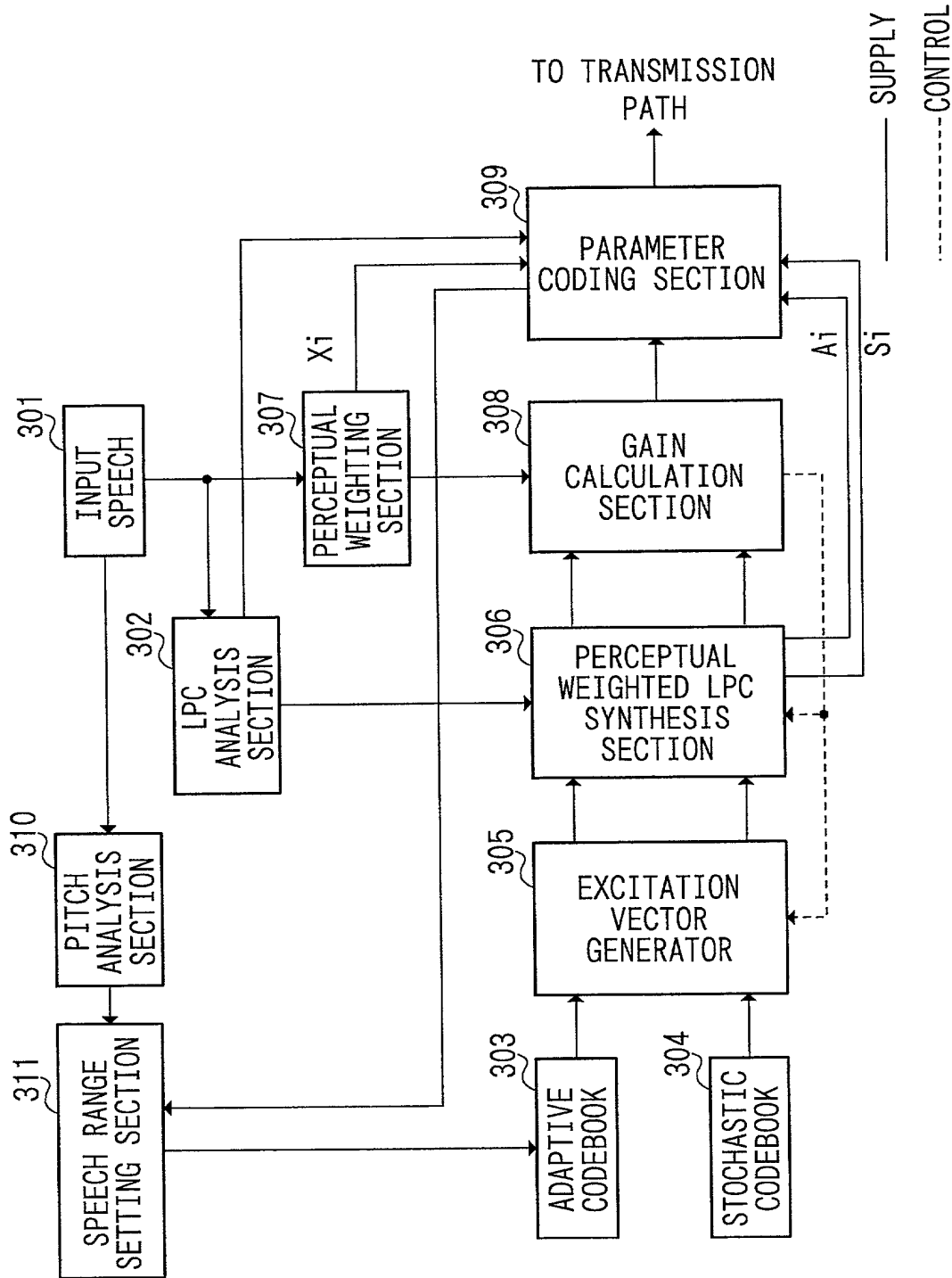


FIG. 7

8 / 1 4

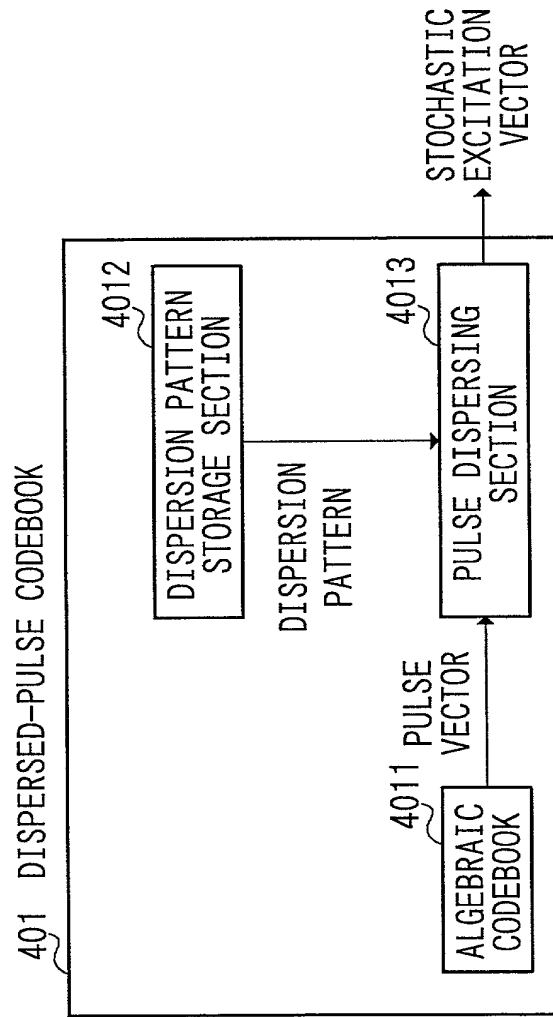


FIG. 8

9 / 1 4

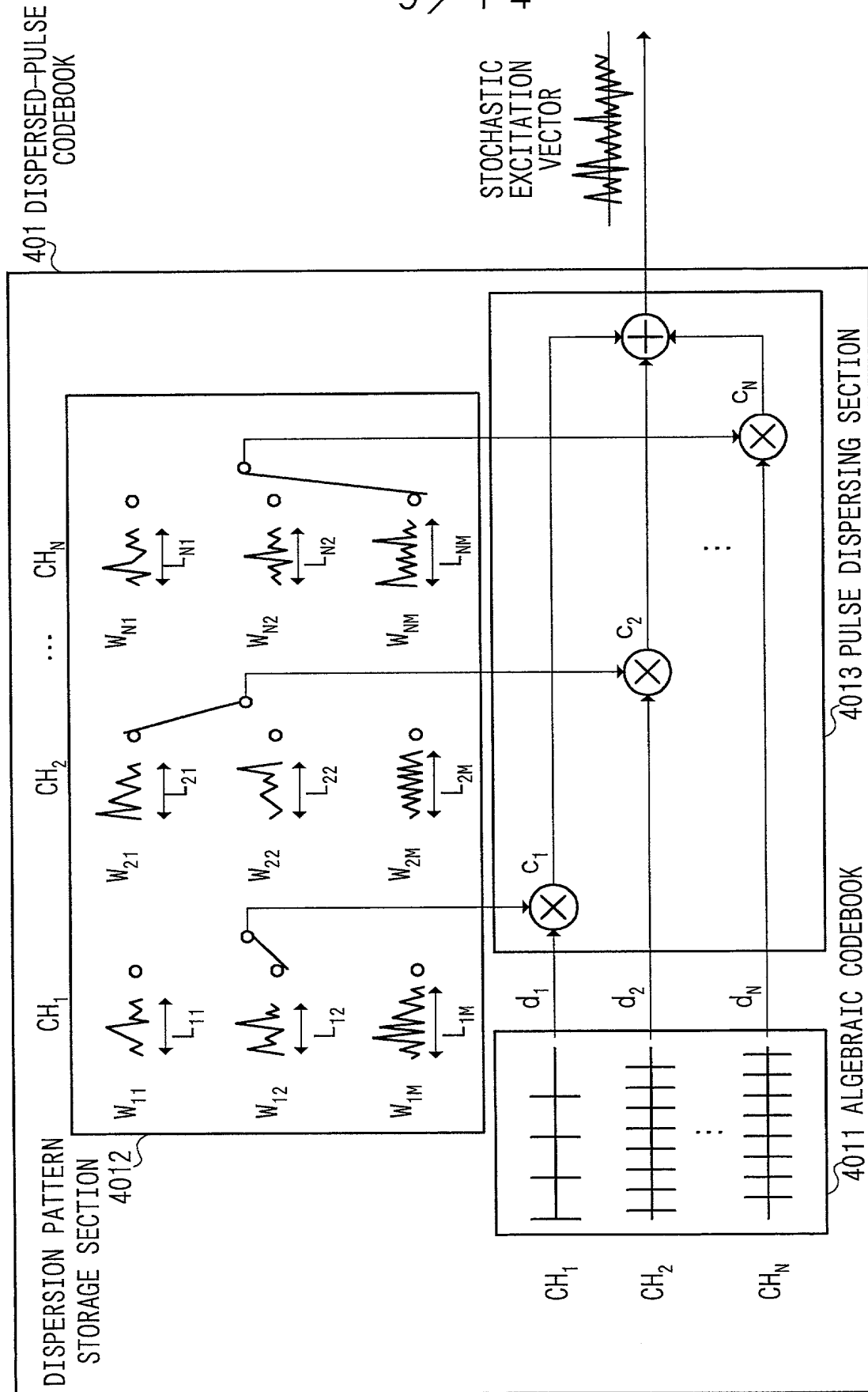


FIG. 9

10/14

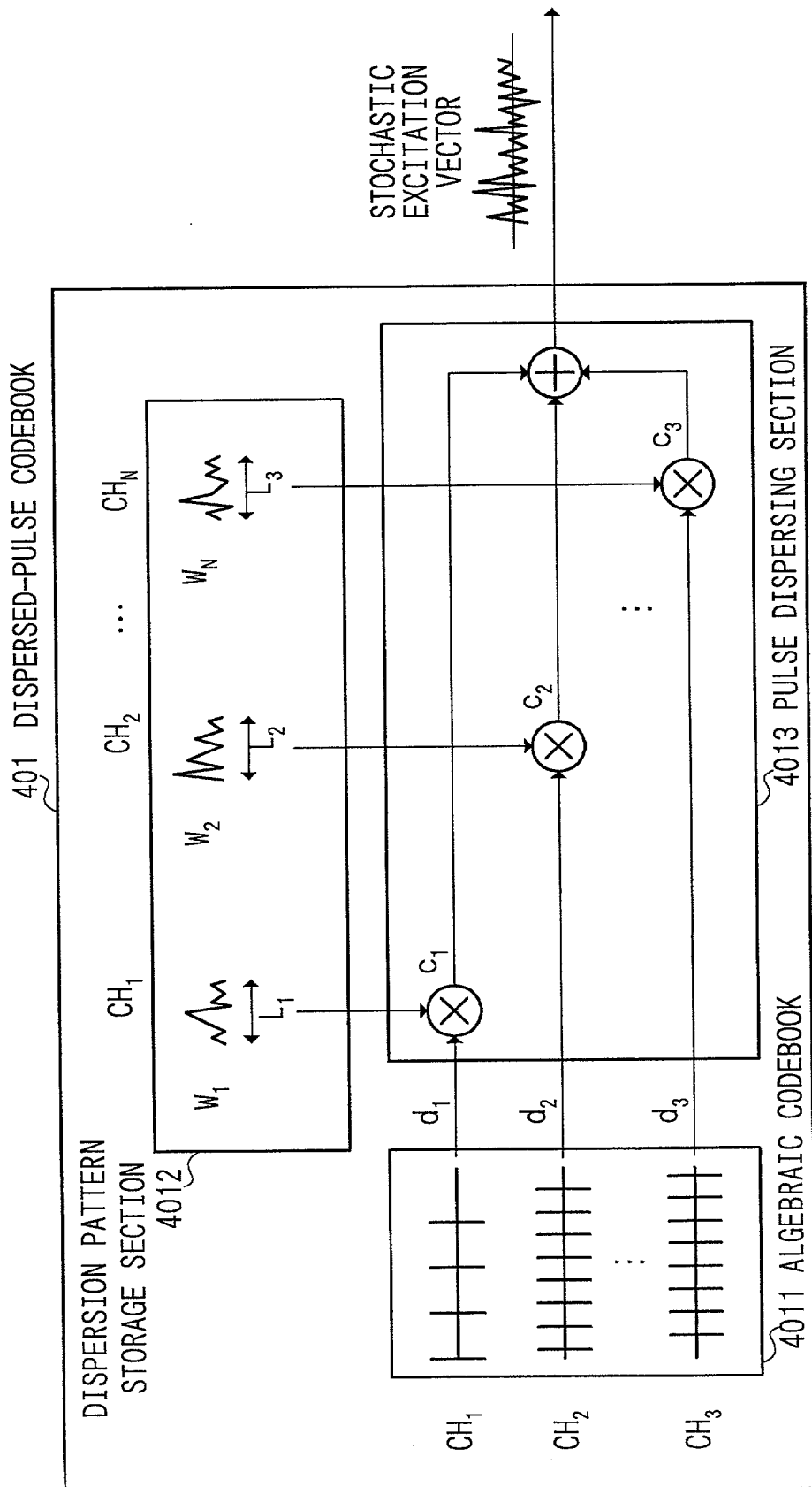


FIG. 10

11 / 14

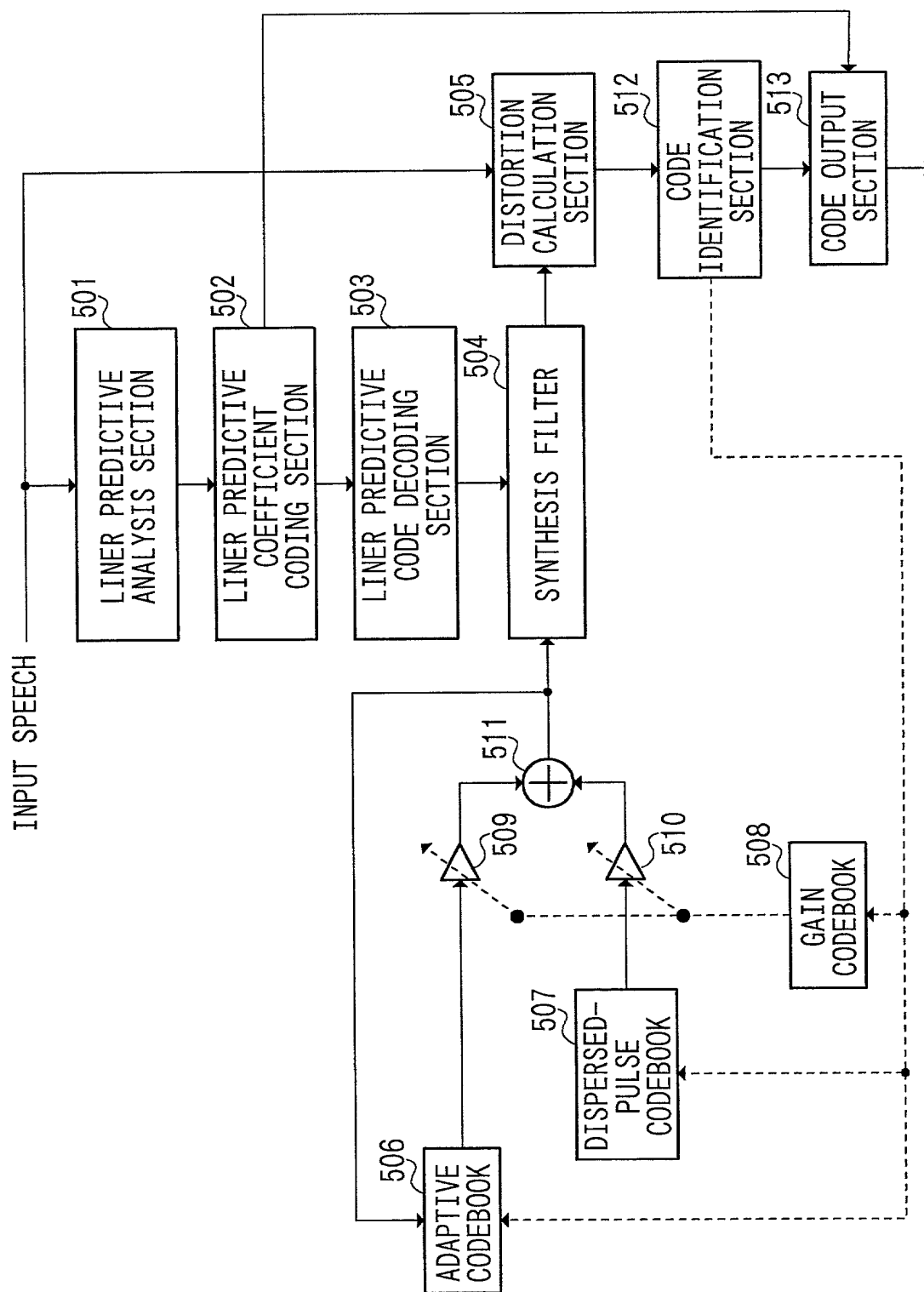


FIG. 11

1 2 / 1 4

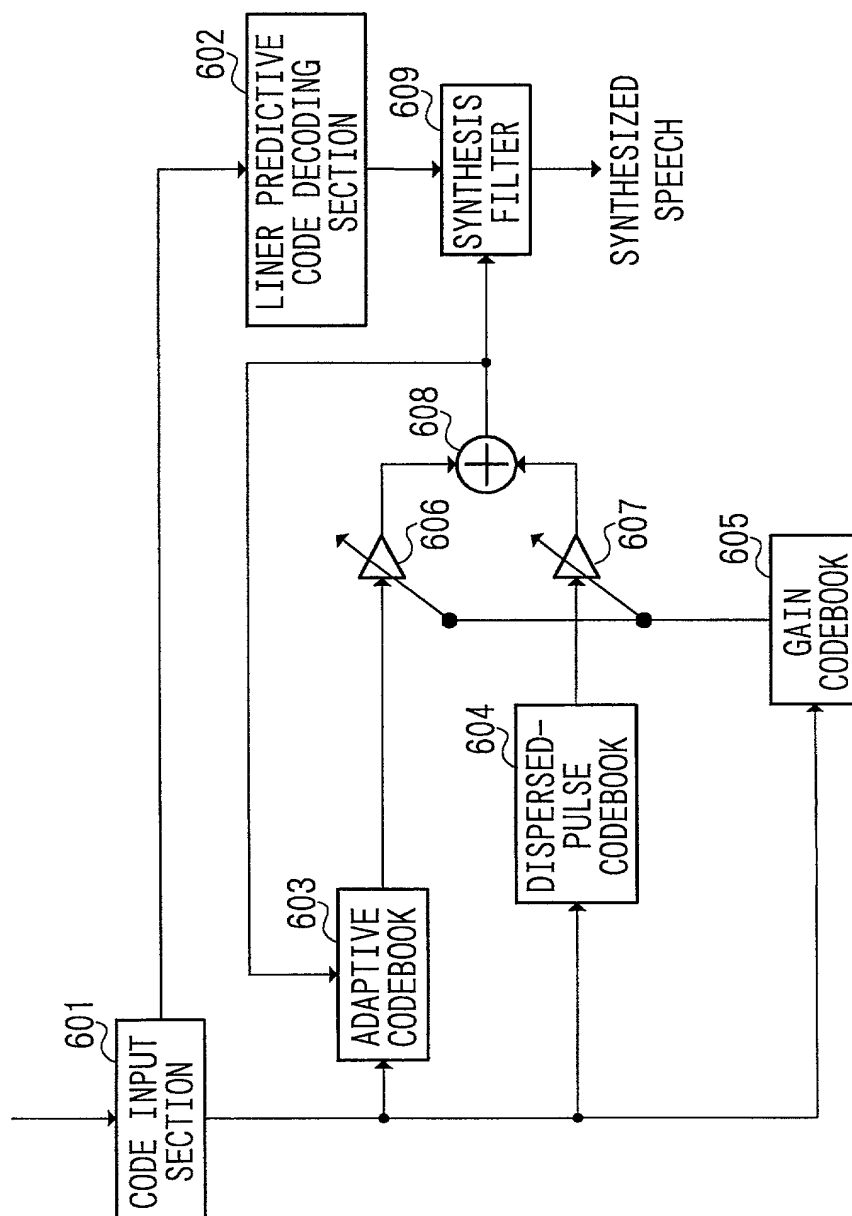


FIG. 12



13/14

## 507 DISPERSED-PULSE CODEBOOK

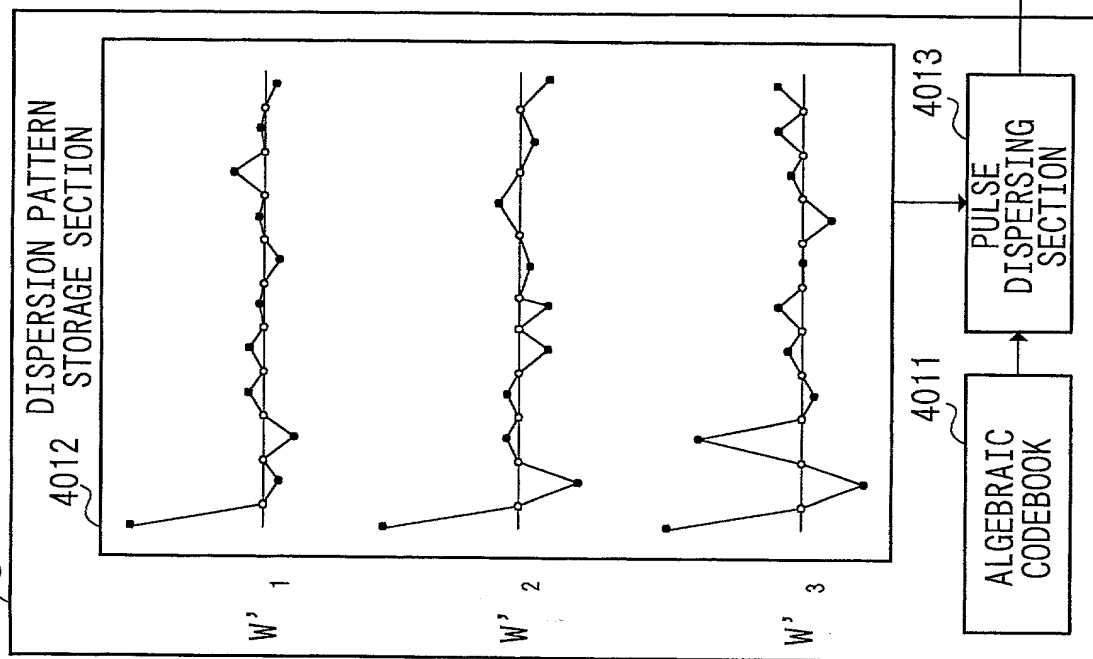


FIG. 13A

## 604 DISPERSED-PULSE CODEBOOK

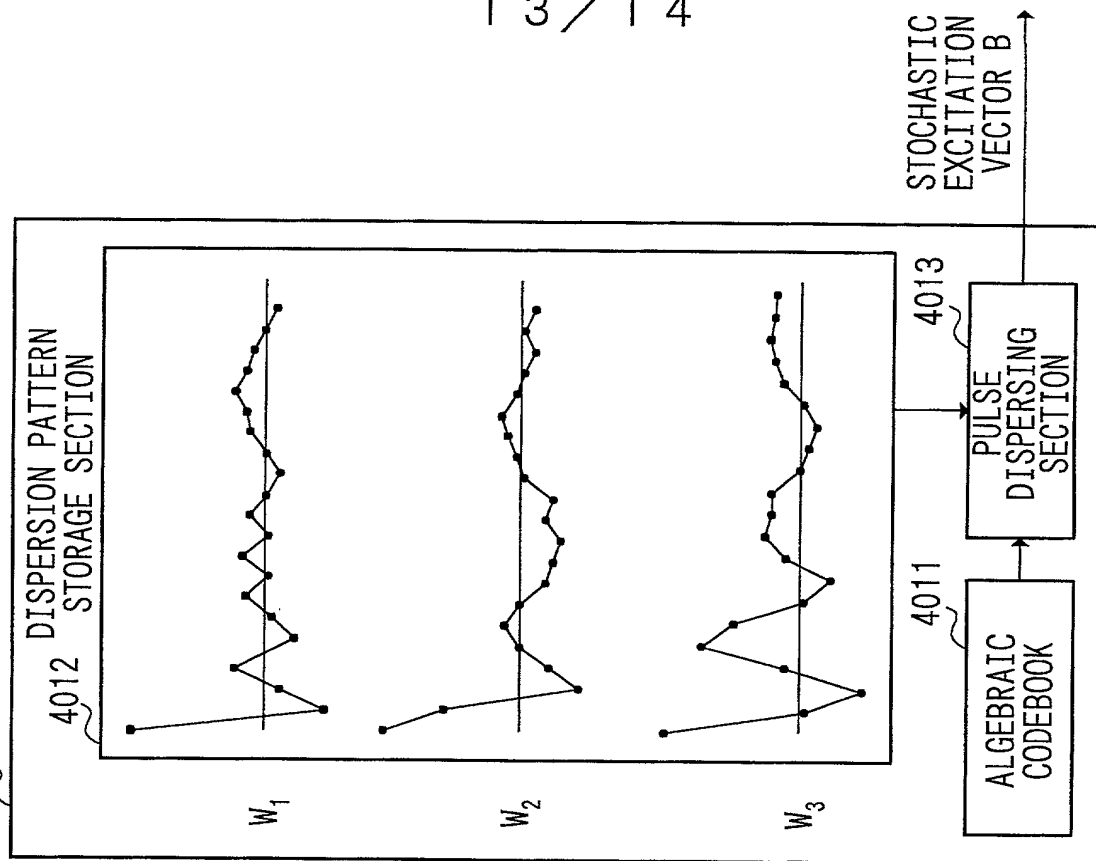


FIG. 13B

1 4 / 1 4

604 DISPERSED-PULSE CODEBOOK

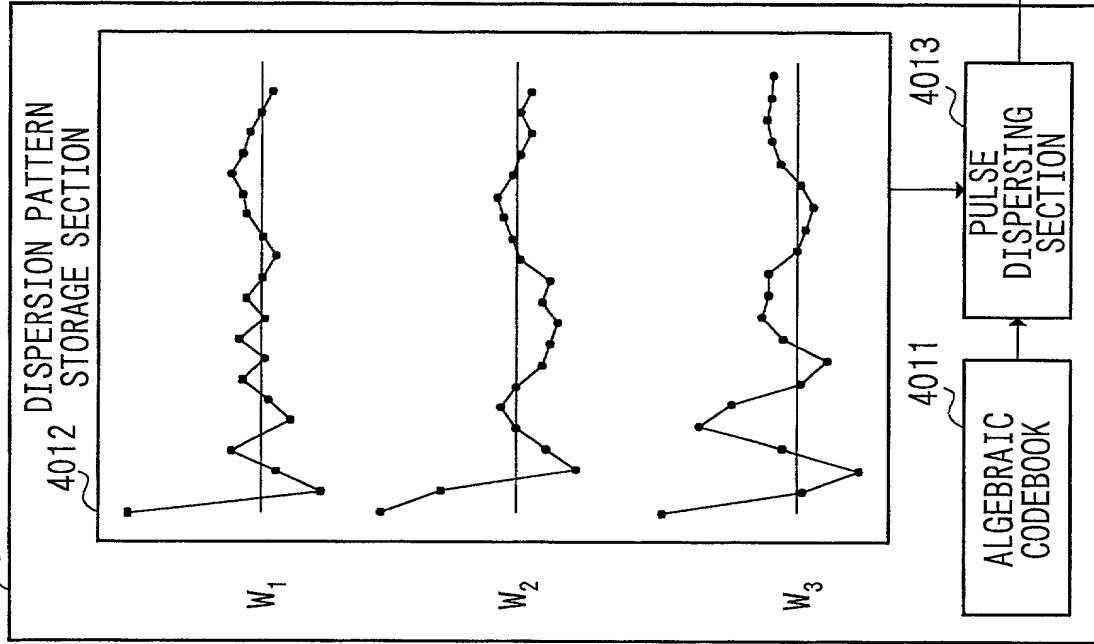


FIG. 14B

507 DISPERSED-PULSE CODEBOOK

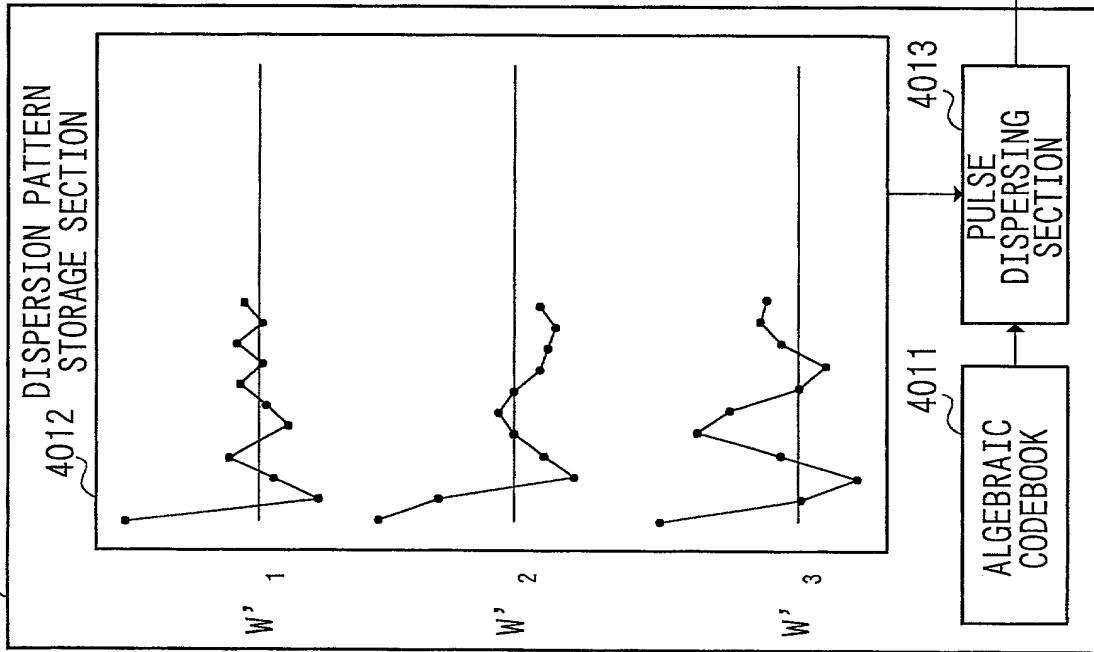


FIG. 14A

# Declaration and Power of Attorney For Utility or Design Patent Application

## 特許出願宣言書

### Japanese Language Declaration

私は、下欄に氏名を記載した発明者として、以下のとおり  
宣言する:

私の住所、郵便の宛先および国籍は、下欄に氏名に続いて記載したとおり  
であり、

名称の発明に関し、請求の範囲に記載した特許を求める主題の本来の、  
最初にして唯一の発明者である(一人の氏名のみが下欄に記載されている  
場合)か、もしくは本来の、最初にして共同の発明者である(複数の氏名が  
下欄に記載されている場合)と信じ、

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

上記発明の明細書(下記の欄でX印がついていない場合は、  
本書に添付)は、

☐ 年 月 日に提出され、  
米国出願番号 \_\_\_\_\_ とし、  
(該当する場合) 年 月 日に訂正されました。又は、

特許協定条約国際出願番号 \_\_\_\_\_ とし、  
(該当する場合) 年 月 日に訂正されました。

私は、前期のとおりに補正した請求の範囲を含む前記明細書の内容を検  
討し、理解したことを陳述する。

私は、連邦規則法典第37編第1条第56項定義されるとおり、特許資格  
の有無について重要な情報を開示すべき義務があることを認めます。

私は合衆国法典第35部第119条(a-d)項又は第365条(b)項に基づく、  
下記の外国特許出願又は発明者証出願、或いは第365条(a)項に基づく、  
少なくとも米国以外の1ヶ国を指名したPCT国際出願の外国優先権を主張  
し、更に優先権の主張に係わる基礎出願の出願日前の出願日を有する外国  
特許出願、又は発明者証出願或いはPCT国際出願を以下に"なし"の箱に  
印をつけることにより明記する:

Prior foreign applications  
先の外国出願

JP11-235050

(Number)  
(番号)

JAPAN

(Country)  
(国名)

23/August/1999

(Day/Month/Year Filed)  
(出願の年月日)

JP11-236728

(Number)  
(番号)

JAPAN

(Country)  
(国名)

24/August/1999

(Day/Month/Year Filed)  
(出願の年月日)

JP11-248363

(Number)  
(番号)

JAPAN

(Country)  
(国名)

02/September/1999

(Day/Month/Year Filed)  
(出願の年月日)

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated  
below next to my name.

I believe I am the original, first and sole inventor (if only one name  
is listed below) or an original, first and joint inventor (if plural  
names are listed below) of the subject matter which is claimed  
and for which a patent is sought on the invention entitled

### APPARATUS AND METHOD FOR SPEECH CODING

\_\_\_\_\_  
\_\_\_\_\_

the specification of which is attached hereto unless the following  
box is checked:

☐ was filed on \_\_\_\_\_ as  
United States Application Number \_\_\_\_\_  
and was amended on \_\_\_\_\_ (if applicable) or,

PCT International Application Number PCT/JP99/03064  
and was amended on \_\_\_\_\_ (if applicable)

I hereby state that I have reviewed and understand the contents of  
the above identified specification, including the claims, as  
amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to  
patentability as defined in Title 37, Code of Federal Regulations,  
§1.56.

I hereby claim foreign priority under Title 35, United States Code  
§119(a-d) or §365(b) of any foreign application(s) for patent or  
inventor's certificate, or §365(a) of any PCT international  
application which designated at least one country other than the  
United States, listed below. I have also identified below, by  
checking the "No" box, any foreign application for patent or  
inventor's certificate, or of any PCT international application having  
a filing date before that of the application on which priority is  
claimed:

Priority claimed  
優先権の主張

☒ ☐

Yes No  
あり なし

☒ ☐

Yes No  
あり なし

☒ ☐

Yes No  
あり なし

☐ その他の外国特許出願番号は別紙の追補優先権欄にて記載する。

☐ Additional foreign application numbers are listed on a  
supplemental priority sheet attached hereto.

# Japanese Language Utility or Design Patent Application Declaration

私は、合衆国法典第35部第119条(e)項に基づく、下記の合衆国仮特許出願の利益を主張する。

I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

(Application Number)  
(番号)

(Day/Month/Year Filed)  
出願の年月日

(Application Number)  
(番号)

(Day/Month/Year Filed)  
出願の年月日

(Application Number)  
(番号)

(Day/Month/Year Filed)  
出願の年月日

☐ その他の合衆国仮特許出願番号は別紙の追補優先権欄にて記載する。

☐ Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

私は、合衆国法典第35部第120条に基づく下記の合衆国特許出願、又は第365条(c)項に基づく合衆国を指名したPCT国際出願の利益を主張し、本願の請求の範囲各項に記載の主題が合衆国法典第35部第112条第1項規定の態様で、先の合衆国特許出願又はPCT国際出願に開示されていない限度において、先の出願の出願日と本願の国内出願日又はPCT国際出願日の間に有効となった連邦規則法典第37部第1章第56条に記載の特許要件に所要の情報を開示すべき義務を有することを認める。

I hereby claim the benefit under Title 35, United States Code §120 of any United States application(s), or §365(c) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

(Application No.)  
(出願番号)

(Day/Month/Year Filed)  
(出願の年月日)

(現況)  
(特許済み、係属中、放棄済み)

(Status)  
(patented, pending, abandoned)

(Application No.)  
(出願番号)

(Day/Month/Year Filed)  
(出願の年月日)

(現況)  
(特許済み、係属中、放棄済み)

(Status)  
(patented, pending, abandoned)

☐ その他の合衆国又は国際特許出願番号は別紙の追補優先権欄にて記載する。

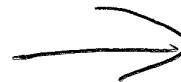
☐ Additional U. S. or international application numbers are listed on a supplemental priority sheet attached hereto.

私は、ここに自己の知識にもとづいて行った陳述がすべて真実であり、自己の有する情報および信ずるところに従って行った陳述が真実であると信じ、さらに故意に虚偽の陳述等を行った場合、合衆国法典第18部第1001条により、罰金もしくは禁錮に処せられるか、またはこれらの刑が併科され、またかかる故意による虚偽による陳述が本願ないし本願に対して付与される特許の有効性を損なうことがあることを認識して、以下の陳述を行ったことを宣言する。

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

私、下記署名者は、ここに記載の米国弁護士または代理人に本出願に関し特許商標庁にて取られるいかなる行為に関して、同米国弁護士又は代理人が、私に直接連絡なしに私の外国弁護士或いは法人代表者からの指示を受け取り、それに従うようここに委任する。この指示を出す者が変更の場合には、ここに記載の米国弁護士又は代理人にその旨通知される。

The undersigned hereby authorizes the U.S. attorney or agent named herein to accept and follow instructions from either his foreign patent agent or corporate representative, if any, as to any action to be taken in the Patent and Trademark Office regarding this application without direct communication between the U.S. attorney or agent and the undersigned. In the event of a change in the persons from whom instructions may be taken, the U.S. attorney or agent named herein will be so notified by the undersigned.



## Japanese Language Utility or Design Patent Application Declaration

委任状： 私は、下記発明者として、下記に明記された顧客番号を伴う以下の弁護士又は、代理人をここに選任し、本順の手続きを遂行すること並びにこれに関する一切の行為を特許商標庁に対して行うことを委任する。そして全ての通信はこの顧客番号宛に発送される。

顧客番号 7055

現在選任された弁護士は下記の通りである。

Neil F. Greenblum  
Bruce H. Bernstein  
James L. Rowland  
Arnold Turk

POWER OF ATTORNEY: As a named inventor, I hereby appoint the attorney(s) and/or agent(s) associated with the Customer Number provided below to prosecute this application and transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to that Customer Number:

**CUSTOMER NUMBER 7055**

The appointed attorneys presently include:

Reg. No. 28,394  
Reg. No. 29,027  
Reg. No. 32,674  
Reg. No. 33,094

Address: **GREENBLUM & BERNSTEIN, P.L.C.**  
1941 ROLAND CLARKE PLACE  
RESTON, VA 20191

直接電話連絡先 : (名称および電話番号)

Direct Telephone Calls to: (name and telephone number)

**GREENBLUM & BERNSTEIN, P.L.C.**  
(703)716-1191

唯一のまたは第一の発明者の氏名	1-00	Full name of sole or first inventor	Kazutoshi YASUNAGA
同発明者の署名	日付	Inventor's signature	Date Kazutoshi Yasunaga February 8, 2001
住所		Residence	Kawasaki-shi, Kanagawa, Japan JPX
国籍		Citizenship	Japan
郵便の宛先		Post Office Address	3-33-17-305, Sugao, Miyamae-ku, Kawasaki-shi, Kanagawa 216-0015 Japan
第二の共同発明者の氏名(該当する場合)	2-00	Full name of second joint inventor, if any	Toshiyuki MORII
同第二共同発明者の署名	日付	Second Inventor's signature	Date Toshiyuki Morii February 8, 2001
住所		Residence	Kawasaki-shi, Kanagawa Japan JPX
国籍		Citizenship	Japan
郵便の宛先		Post Office Address	2-3-7-501, Nijigaoka, Asao-ku, Kawasaki-shi, Kanagawa 215-0015 Japan

(第三またはそれ以降の共同発明者に対しても同様な情報および署名を提供すること。)

(Supply similar information and signature for third and subsequent joint inventors.)